

```

TTCAGGAATT GATTGTAAGT TCTGTTATCA ACTAAAGCCG ACAGCGGTTG      50
CTGATATGCT CTAGAACAAC ATTGCTAAGG CTCACGGTGG TTACTCCGTG      100
TTCAGTGGTG TCGGTGAGCG TACCCGTGAG GGTAACGATT TGTACCATGA      150
AATGCAAGAG ACCCGTGTCA TTCAACTCGA CGGAGAGTCC AAGGTCGCTC      200
5  TTGTCTTCGG TCAAATGAAC GAGCCCCCTG GTGCCCCGTGC CCGTGTTGCC      250
CTTACCGGTT TGACCATTGC TGAATACTTC CGTGACGAGG AAGGCCAAGA      300
CGGTAGGCTT CATGCTTCTA TCGCTAGGGG CGTGTGATAC AGGAGGCTAA      350
TCGCTTTTCT AGTGCTTCTC TTTATTGACA ACATTTTCCG TTTCAC TCAA      400
GCTGGTTCTG AAGTGTCTGC CTTGCTCGGT CGTATTCCTT CCGCTGTCGG      450
10 TTACCAACCT ACTCTCGCCG TCGACATGGG TGTATGCAG GAACGTATCA      500
CCACCACCAC CAAGGGATCC ATTACTTCAG TGCA                        534

```

15 2) INFORMATION FOR SEQ ID NO: 480

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 494 bases
(B) TYPE: Nucleic acid
20 (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

25 (vi) ORIGINAL SOURCE:

```

(A) ORGANISM: Cryptococcus albidus
(B) STRAIN: ATCC 66030

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 480

```

30 CGTCTTGATT CAAGAATTGA TCAACAACAT CGCCAAGGCC CACGGTGGTT      50
ACTCGGTCTT CACCGGTGTC GGTGAGCGAA CTCGAGAGGG TAACGATCTG      100
TACCACGAAG TGAGTTGCGC CGTCCGAGTG TTTCCCGGGG AATCGCAAGA      150
CTGATGTTGT CCCTTCTTCT CAGATGAGGG AAACCGGTGT CATCAACCTC      200
35 GAGGGTGACT CCAAGGTCGC CTTGGTCTTC GGTCAGATGA ACGAGCCCCC      250
TGGAGCCCCG GCCCGAGTCG CCTTGACCGG TTTGACCATT GCCGAATACT      300
TCCGAGACGA GGAGGGTCAG GATGTCTTGT TGTTTCATTGA CAACATTTTC      350
CGATTACCC AAGCCGGTTC CGAAGTGTCC GCCTTGTTGG GTCGTATCCC      400
CTCCGCCGTC GGTTACCAGC CCACTCTGTC CACCGACATG GGTACCATGC      450
40 AGGAGCGAAT TACCACCACC AAGAAGGGTT CCATCACTTC CGTC          494

```

45 2) INFORMATION FOR SEQ ID NO: 481

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 415 bases
(B) TYPE: Nucleic acid
50 (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

```

55 (A) ORGANISM: Fusarium oxysporum
(B) STRAIN: WSA-212

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 481

```

60 TTCAGGAGCT TATCAACAAC ATCGCCAAGG CCCACGGTGG TTACTCCGTC      50

```

	TTCACTGGTG	TCGGTGAGCG	AACTCGTGAG	GGTAACGATC	TGTACCACGA	100
	AATGCAGGAG	ACTTCCGTCA	TTCAGCTTGA	TGGCGAGTCC	AAGGTCGCCC	150
	TGGTCTTCGG	TCAGATGAAC	GAGCCCCCTG	GAGCTCGTGC	CCGTGTCGCC	200
	CTTACCGGTC	TTACTGTAGC	TGAATACTTC	AGAGATGAGG	AGGGTCAGGA	250
5	CGTGCTGCTC	TTCATTGACA	ACATTTTCCG	ATTCACTCAG	GCCGGTTCCG	300
	AGGTGCTGTC	CCTTCTCGGT	CGTATCCCCT	CTGCCGTCGG	TTACCAGCCC	350
	ACCCTCGCCG	TCGACATGGG	TGGTATGCAA	GAGCGTATTA	CCACCACCAC	400
	CAAGGGTTCC	ATTAC				415

10

## 2) INFORMATION FOR SEQ ID NO: 482

## (i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 1281 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Geotrichum* spp.  
(B) STRAIN: Lev-4

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 482

	AGGACAACCT	CCCCGCTATT	CTTAACGCTC	TTGAGCTTAA	GAGAGATAAC	50
	GGTGAGAAGC	TCGTTCTCGA	GGTTGCCAG	CATCTGGGTG	AGAACACTGT	100
30	CCGTACTATT	GCTATGGACG	GTACTGAGGG	TCTCGTCCGT	GGTCAGCCCCG	150
	TTGTTGACAC	CGGTGCCCCC	ATTACCATT	CCGTTGGTCG	TGGTACTCTT	200
	GGTAGAATTA	TCAACGTCAT	TGGTGAGCCC	ATCGATGAGC	GTGGACCCAT	250
	TGAGGCTACC	AAGTACCTCC	CCATCCACAC	CGAGCCCCC	ACCTTCGCTG	300
	AGCAGTCTAC	CTCCGCTGAG	GTTCTTGAGA	CTGGTATCAA	GGTTGTCGAT	350
35	CTCCTTGCCC	CCTACGCCCC	TGGTGGTAAG	ATTGGTCTCT	TCGGTGGTGC	400
	CGGTGTCGGT	AAGACCGTTT	TCATTCAGGA	GCTGATTAA	AACATTGCCA	450
	AGGCCCATGG	TGGTTTCTCC	GTTTTACCG	GTGTCGGTGA	GAGAACCCGT	500
	GAGGGTAACG	ATTTGTACCG	TGAGATGAAG	GAGACCGGTG	TCATCAACCT	550
	CGAGGGTGAG	TCTAAGGTCG	CTCTCGTTTT	CGGTCAGATG	AACGAGCCCC	600
40	CTGGAGCCCC	TGCCCGTGTT	GCTCTTACTG	GTCTTACCAT	TGCTGAGTAC	650
	TTCCGTGATG	AGGAGGGTCA	GGATGTGTTG	CTCTTCGTTG	ACAACATTTT	700
	CCGTTTCACT	CAGGCCGGTT	CCGAGGTGTC	TGCCCTTTTG	GGTCGTATTC	750
	CCTCCGCTGT	CGGTTACCAG	CCCACCCTTG	CCACTGATAT	GGGTGCCCTG	800
	CAAGAGCGTA	TTACCACCAC	CCAGAAGGGT	TCCGTCACTT	CCGTCCAGGC	850
45	CGTCTACGTC	CCTGCCGATG	ATTTGACCGA	TCCTGCCCCC	GCTACCACCT	900
	TCGCCCATT	GGATGCCACC	ACCGTCTTGT	CGCGTTCCAT	TTCTGAGTTG	950
	GGTATCTACC	CCGCTGTCGA	TCCCCTTGAT	TCCAAGTCTC	GTCTTTTGGA	1000
	TATCACCGTC	GTTGGCCAGG	AGCACTACGA	TGTTGCTACC	CAGGTCCAGC	1050
	AGACCCTCCA	GTCTTACAAG	TCTCTTCAGG	ATATCATTCG	CATTTTGGGT	1100
50	ATGGATGAGT	TGTCTGAGGC	TGATAAGCTT	ACTGTCGAGC	GTGCCCCGTA	1150
	GATCCAGAGA	TTCTTTTCCC	AGCCCTTCAC	TGTCGCTGAG	GTTTTCACTG	1200
	GTATCGAGGG	CCGTCTCGTT	CCTTTGAAGG	ACACTGTTCG	CTCTTTCAAG	1250
	GAGATCCTTG	AGGGCAAGTA	CGACCACCTC	C		1281

55

## 2) INFORMATION FOR SEQ ID NO: 483

## (i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 586 bases



(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii)MOLECULE TYPE: Genomic DNA

(vi)ORIGINAL SOURCE:

(A) ORGANISM: *Histoplasma capsulatum*  
 (B) STRAIN: G185A5

10

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 483

ATTCAAGAAT	TGATCGTACG	TTCCTCCGCC	CCACACACGA	TCAATGGAGA	50
AAGAAACAAA	TTTTTTGGGA	GTGGTCAATT	TTTCTAATAA	TTCGAATAGA	100
15 ACAACATTGC	CAAAGCCCAC	GGTGGTTACT	CCGTGTTTAC	TGGTGTCCGC	150
GAGCGGACCC	GTGAAGGAAA	TGACTTGTA	CACGAAATGC	AGGAAACCCG	200
TGTTATCCAG	CTCGATGGAG	AGTCCAAGGT	CGCCCTCGTT	TTCGGTCAGA	250
TGAACGAGCC	TCCCGGAGCC	CGTGCCCGTG	TTGCCCTCAC	TGGTCTGACC	300
GTTGCCGAGT	ACTTCCGTGA	CGAGGAAGGC	CAAGACGGTA	TGTATAAGTA	350
20 TACACCGTAG	CAAATCAACA	CAGAGCTTCA	CTCACGCTCG	GATTTAGTGC	400
TTCTCTTCAT	CGACAACATT	TTCCGCTTCA	CTCAGGCCGG	TTCCGAAGTG	450
TCTGCCCTGC	TCGGCCGTAT	TCCCTCCGCC	GTCCGTTACC	AACCCACCCT	500
CGCCGTGGAC	ATGGGTGGTA	TGCAGGAACG	TATCACCAC	ACCACCAAGG	550
25 GCTCCATCAC	CTCTGTGCAR	GCCGTCTACG	TCCCCG		586

2)INFORMATION FOR SEQ ID NO: 484

30 (i)SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1145 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

35

(ii)MOLECULE TYPE: Genomic DNA

(vi)ORIGINAL SOURCE:

(A) ORGANISM: *Malassezia furfur*  
 (B) STRAIN: ATCC 42132

40

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 484

GCGTGGCCAG	AAGGTCATTG	ACACTGGTGC	TCCCATCACC	ATCCCCGTCG	50
45 GTGGTGCCAC	ACTGGGTCGT	ATTCTGAACG	TCACGGGTGA	CCCTATTGAC	100
GAGCGTGGCC	CCGTTAAGAC	TGACGTTTTT	CGCCCCATTC	ACCGTGACCC	150
CCCTGCCTTT	GTCGAGCAGT	CGACTGATGC	CGAGATTCTC	GAGACTGGTA	200
TCAAAGTCGT	TGACCTGATT	GCCCCCTTAC	CCCGTGGTGG	TAAGATTGGT	250
CTGTTTCGGTG	GTGCCGGTGT	CGGTAAGACC	GTGCTTATCC	AGGAGCTCAT	300
50 CAACAACATC	GCCAAGGCCC	ACGGTGGTTT	CTCCGTGTTC	ACTGGTGTCTG	350
GTGAGCGTAC	TCGTGAGGGT	AACGATTTGT	ACCACGAGAT	GATTGAAACC	400
GGTGTCATTA	ACCTCGAGGG	TGACTCGAAG	GTGGCTCTGG	TGTTTCGGTCA	450
GATGAACGAG	CCCCCGGGTG	CCCGTGCGCG	TGTCGCTCTT	ACTGGTCTGA	500
CTGTGGCCGA	GTAATTCCGT	GACGACGAG	GCCAGGATGT	GCTGCTGTTC	550
55 ATTGACAACA	TTTTCCGTTT	CACCCAGGCC	GGTTCGGAGA	CTTCGGCTCT	600
GCTGGGTCGT	ATCCCTTCGT	CGGTTCGGTTA	CCAGCCCACT	TTGTTCGACCG	650
ATATGGGTGC	CATGCAGGAG	CGTATCACCA	CCACCAAGAA	GGGTTCGATT	700
ACGTCGGTGC	AGGCCGTCTA	CGTGCCCGCC	GATGATGTCA	CTGACCCTGC	750
CCCTGCCACT	ACCTTCGCCC	ACCTTGACGC	TACCACTGTG	CTTGACCGTT	800
60 CGATCGCTGA	GCTGGGTATC	TACCCCGCTG	TTGACCCGTT	GAACCTCGAAC	850

```

TCGCGTATGC TTGACCCCGC TATTGTGGGT CAGGAGCACT ACGACGTGGC 900
CTCTGGCGTG CAGAAGCTGC TCCAGGACTA CAAGTCGCTT CAAGATATCA 950
TTGCCATTCT GGGTATGGAT GAGCTTTCTG AGGAGGACAA GCTCACTGTC 1000
GAGCGTGCCC GTAAGATGCA GCGTTTCATG TCGCAGCCTT TCGCTGTCGC 1050
5 CCAGGTCTTT ACTGGTATCG AAGGTCGTCT TGTGGCCCTG AAGGACACGA 1100
TCAAGGCCTG CAAGGAGATC CTGTCGGGCA AGCACGACAA CCTCC 1145

```

10 2) INFORMATION FOR SEQ ID NO: 485

(i) SEQUENCE CHARACTERISTICS:

```

15 (A) LENGTH: 1261 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Malassezia pachydermatis
    (B) STRAIN: ATCC 42756

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485

```

25 TCCCCGCCATC TTCAACGCCC TGGAGGTCCA GGACATGAAG AACGGTGGCC 50
   GCCTTGTTCT GGAGGTTGCC CAGCACCTTG GTGAGAACAC TGTTCGTTGC 100
   ATTGCTATGG ACGGTACCGA GGGTCTTGTC CGTGGTCAGA AGGTCCTTGA 150
   CACTGGTGCC CCGATCACTA TCCCTGTCGG TAACGGTACC TTGGGCCGYA 200
30 TCCTGAACGT CACTGGTGAG CCTGTGGATG AGCGTGGTCC GGTAAAGACT 250
   GACGTCTACC GTCCAATCCA CCGTGAGCCC CCGGCGTTCC TTGAGCAGTC 300
   GACTGATGCT GAGATTCTTG AGACTGGTAT CAAGGTGGTC GACCTGCTCG 350
   CCCCATACGC TCGTGGTGGT AAGATTGGTC TWTTCGGTGG TGCYGGTGTY 400
   GGTAAGACCG TGCTGATTCA GGAGCTTATC AACAACTTG CCAAGGCCCA 450
35 CCGTGGTTTC TCGGTGTTCA CTGGTGTCCG TGAGCGTACT CGTGAGGGTA 500
   ACGATCTGTA CCATGAAATG ATTGAGACTG GTGTCATCAA CGTTGACGGT 550
   GACTCGAAGG TCGCTCTCGT GTTCGGTCAG ATGAACGAGC CCCCAGGTGC 600
   CCGTGCCCGT GTCGCCCTGA CYGGTCTGAC CATCGCCGAG TACTTCCGTG 650
   ACGACGAGGG TCAGGATGTG CTGCTCTTCA TTGACAACAT TTTCCGTTTC 700
40 ACTCAGGCTG GTTCGGAGAC TTCGGCTCTG CTGGGTCGTA TCCCGTCGGC 750
   TGTCCGGTTAC CAGCCACGCC TTGCCACGGA TATGGGTGCC ATGCAGGARC 800
   GTATYACCAC CACCAAGAAG GGTTCGATTA CCTCGGTGCA GGCYGTTTAC 850
   GTTCCGGCCG ACGATGTGAC TGACCCTGCC CCGGCCACGA CCTTCGCCCA 900
   CTTGGACGCC ACCACGGTGT TGGACCGTTC GATTGCTGAG CTGGGTATCT 950
45 ACCCGGCCGT CGACCCGCTG AACTCGAAGT CGCGTATGCT TGACCCGTCA 1000
   ATTGTCCGGT TGGAGCACTA CAACGTTGCT TCGGGTGTCC AGAAGCTTCT 1050
   YCAGGACTAC AAGTCGCTCC AAGATATCAT TGCCATTCTG GGTATGGATG 1100
   AGTTGTGCGA GGAGGACAAG CTCACTGTCT AGCGTGCCCC TAAGATGCAG 1150
   CGTTTCCTGT CGCAGCCTTT CGCTGTGGCC CAGGTCTTCA CTGGTATCGA 1200
50 GGGTCGTCTT GTGTCGCTCA AGGACACCAT CAACGCCTGC AAGGAGATTC 1250
   TGTCCGGTAA G 1261

```

55 2) INFORMATION FOR SEQ ID NO: 486

(i) SEQUENCE CHARACTERISTICS:

```

60 (A) LENGTH: 1282 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double

```

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Metschnikowia pulcherrima*

(B) STRAIN: DSM 70336

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 486

10 AGGAGGGCAA CTTGCCAGCT ATCTTCAACG CTTTGACGTT GAAGAACGGC 50  
 GACCAGAAGT TGGTCTTGGA GGTGGCCAG CACTTGGGTG AGAACACCGT 100  
 CAGAACCATT GCCATGGACG GTACCGAGGG TTTGGTCAGA GGCGCCTCTG 150  
 TCACCGACAC YGGTGCCCTT ATCTCCGTGC CTGTCGGCCG TGAGACCTTG 200  
 15 GGTTCGTATTA TCAATGTTGT TGGTGAGCCA ATCGACGAGA GAGGCCCAAT 250  
 CAACACCAAG AAGAGAAACC CTATTCACAC CGACCCACCT TCGTTTGTCC 300  
 AGCAATCCAC TTCCGCCGAG GTCTTGGAGA CTGGTATCAA GGTTGTCGAC 350  
 TTGTTGGCCC CTTACGCCAG AGGTGGTAAG ATTGGTTTGT TCGGTGGTGC 400  
 CGGTGTCGGT AAGACCGTGT TCATCCAGGA GTTGATTAAC AACATTGCCA 450  
 20 AGGCCACGG TGGTTTCTCC GTTTTCACCG GTGTCGGTGA GAGAACCAGA 500  
 GAGGGTAACG ATTTGTACCG TGAAATGCAG GAGACTGGTG TCATCAACTT 550  
 CGAGGGTGAC TCCAAGGTCG CCTTGGTTTT CGGTCAGATG AACGAGCCTC 600  
 CAGGAGCTAG AGCYAGAGTT GCCTTGACCG GTTTGACCAT CGCCGAGTAC 650  
 TTCAGAGACG AGGAGGGTCA GGATGTGTTG TTGTTCTGTCG ACAACATTTT 700  
 25 CAGATTCACC CAAGCCGGTT CTGAGGTGTC TGCCTTGTTG GGTCGTATTC 750  
 CATCCGCTGT CGGATACCAG CCAACCTTGG CCACCGATAT GGGTGCCTTG 800  
 CAGGAGAGAA TTACCACCAC CAAGAAGGGT TCCGTCACCT CCGTCCAGGC 850  
 CGTCTACGTG CCAGCCGATG ACTTGACTGA CCCTGCGCCA GCCACCACTT 900  
 TCGCCCACTT GGACGCCACC ACTGTGTTGT CCAGAGGTAT CTCTGAGTTG 950  
 30 GGTATCTACC CCGCTGTCGA CCCCTTGGAC TCCAACCTCA GATTGTTGGA 1000  
 CGCCACCGTT GTTGGCCAGG AGCACTACGA CGTCGCCACC AACGTCCAGC 1050  
 AAACTTTACA AGCTTACAAG TCCTTGACAG ATATCATTGC CATTTTGGGT 1100  
 ATGGATGAGT TGTCCGAGAC CCGACAAGTT GACCCGTCGA GAGAGCCAGA 1150  
 AAGATCCAGA AGTTCTTGTC CCAGCCATTT GCCGTCGCCG AGGTTTTCAC 1200  
 35 CGGTATTGAG GGTAGATTGG TTAGATTGGA GGACACCGTT AGATCCTTTA 1250  
 AGGAGGTTTT GGAAGGTAAG TACGACCACT TG 1282

40 2) INFORMATION FOR SEQ ID NO: 487

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 482 bases

(B) TYPE: Nucleic acid

45 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

50 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Penicillium marneffe*

(B) STRAIN: WSA-214

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 487

55 TGTCTTTATC CAGGAGTTGA TTGTACGTCT TTACCTTTCT GCCTGACTGT 50  
 TTACGACAAC TAACGAAAGC GTAGAACAAC ATTGCCAAGG CTCACGGTGG 100  
 TTA CTCTGTC TTA CTGTTG TCGGTGAACG TACTCGTGAG GGTAACGATT 150  
 TGTACCACGA AATGCAGGAA ACTGGTGTCA TTCAGCTCGA GGGTGAATCC 200  
 60 AAGGTCGCCC TCGTGTTCGG TCAGATGAAC GAGCCCCCG GTGCCCGTGC 250

	CCGTGTCGCT	CTTACTGGTT	TGACCATTGC	CGAGTACTTC	CGTGACGAGG	300
	AAGGTCAGGA	CGTGCTTCTC	TTCATTGACA	ACATTTTCCG	TTTCACTCAG	350
	GCCGGTTCTG	AGGTGCTCTG	CCTTCTGGGT	CGTATCCCCT	CTGCCGTCGG	400
	TTACCAGCCC	ACCCTTGCCG	TCGACATGGG	TATCATGCAG	GAGCGTATTA	450
5	CCACCACCAC	CAAGGGTTCC	ATCACCTCCG	TC		482

## 2) INFORMATION FOR SEQ ID NO: 488

10

## (i) SEQUENCE CHARACTERISTICS:

	(A)	LENGTH: 1290 bases
	(B)	TYPE: Nucleic acid
	(C)	STRANDEDNESS: Double
15	(D)	TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

20	(A)	ORGANISM: <i>Pichia anomala</i>
	(B)	STRAIN: ATCC 18205

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 488

25	TCGAACAAGG	TAACTTACCA	GCTATCTTGA	ATGCTTTAGA	AATCAAAACC	50
	CCAAGTGGTA	CTCCATTAGT	TTTAGAAGTT	GCTCAACATT	TAGGTGAAAA	100
	CACTGTCAGA	ACTATTGCTA	TGGATGGTAC	TGAAGGTTTA	GTTCGTGGTG	150
	AACAAGTTAC	TGATACTGGT	TCTCCAATCA	CTGTCCCAGT	TGGTCGTGAA	200
	ACTTTAGGTC	GTATTATCAA	CGTTGTTGGT	GAACCAATTG	ATGAACGTGG	250
30	TCCAATTAAC	ACCAAAACAAA	GAAACCCAAT	TCACGCTGAA	CCACCTTCAT	300
	TCAGTGAACA	ATCAACTGCT	GCTGAAGTTT	TAGAAACTGG	TATCAAAGTT	350
	GTYGATTTAT	TAGCTCCATA	CGCTAGAGGT	GGTAAAATTG	GTTTATTCGG	400
	TGGTGCCGGT	TGCTGTAAAA	CTGTCTTTAT	CCAAGAATTG	ATTAACAACA	450
	TTGCTAAAGC	TCATGGTGGT	TTCTCAGTTT	TCACCGGTGT	TGGTGAAAGA	500
35	ACCAGAGAAG	GTAACGATTT	ATACCGTGAA	ATGAAAGAAA	CTGGTGTTAT	550
	TAACTTGGA	GGTGATTCTA	AGGTCGCTTT	AGTTTTCGGT	CAAATGAATG	600
	AACCACCAGG	TGCTAGAGCT	CGTGTGCTT	TAACTGGTTT	GACCATTGCT	650
	GAATACTTCA	GAGATGAAGA	AGGTCAAGAT	GTCTTGTTAT	TCGTTGATAA	700
	CATTTTCAGA	TTCAACCAAG	CCGGTTCAGA	AGTTTCTGCC	TTATTAGGTC	750
40	GTATTCCATC	TGCTGTCCGT	TATCAACCAA	CTTTAGCAAC	TGATATGGGT	800
	TTGTTACAAG	AACGTATTAC	CACCACACAA	AAAGGTTTCA	TTACTTCTGT	850
	CCAAGCTGTT	TATGTCCCAG	CTGATGATTT	AACAGATCCT	GCTCCAGCTA	900
	CCACTTTCGC	CCATTTGGAT	GCTACTACTG	TCTTGTCTCG	TGGTATTTCA	950
	GAATTAGGTA	TTTACCCAGC	TGTCGATCCA	TTAGATTCTA	AATCAAGATT	1000
45	ATTAGATGCT	TCAGTTGTTG	GTCAAGAACA	TTATGATGTT	GCTACCAACG	1050
	TTCAACAAAC	TTTACAAGCT	TACAAATCTT	TACAAGATAT	TATTGCTATT	1100
	TTAGGTATGG	ATGAATTGTC	TGAACAAGAT	AAATTGACTG	TCGAAAGAGC	1150
	AAGAAAAATC	CAAAGATTCT	TATCTCAACC	ATTTGCTGTT	GCCGAAGTTT	1200
	TCACTGGTAT	YCCAGGTAGA	TTGGTTAGAT	TAAAAGACAC	TATCAAATCA	1250
50	TTCAAAGATG	TTTTGGAAGG	TAAATATGAT	CACTTACCAG		1290

## 2) INFORMATION FOR SEQ ID NO: 489

55

## (i) SEQUENCE CHARACTERISTICS:

	(A)	LENGTH: 1291 bases
	(B)	TYPE: Nucleic acid
	(C)	STRANDEDNESS: Double
60	(D)	TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Pichia anomala*  
 (B) STRAIN: ATCC 2149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489

```

10 CCAATTCGAA CAAGGTAAC TACCAGCTAT CTTGAATGCT TTAGAAATCA      50
   AAACCCCAAG TGGTACTCCA TTAGTTTTAG AAGTTGCTCA ACATTTAGGT      100
   GAAACACTG TCAGAACTAT TGCTATGGAT GGTACTGAAG GTTTAGTCCG      150
   TGGTGAACAA GTTACTGATA CTGGTTCTCC AATCACTGTC CCAGTTGGTC      200
   GTGAAACTTT AGGTCGTATT ATCAACGTTG TTGGTGAACC AATTGATGAA      250
15 CGTGGTCCAA TTAACACCAA ACAAGAAAC CCAATTCACG CTGAACCACC      300
   TTCATTCACT GAACAATCAA CTGCTGCTGA AGTTTTAGAA ACTGGAATCA      350
   AAGTTGTTGA TTTATTAGCT CCATACGCTA GAGGTGGTAA AATTGGTTTA      400
   TTCGGTGGTG CCGGTGTCGG TAAACTGTC TTTATCCAAG AATTGATTAA      450
   CAACATTGCT AAAGCTCATG GTGGTTTCTC AGTTTTACC GGTGTTGGTG      500
20 AAAGAACCAG AGAAGGTAAC GATTTATACC GTGAAATGAA AGAACTGGT      550
   GTTATTAAC TGGAGGGTGA TTCTAAGGTC GCTTTAGTTT TCGGTCAAAT      600
   GAATGAACCA CCAGGTGCTA GAGCTCGTGT TGCTTTAACT GGTTCGACCA      650
   TTGCTGAATA CTTCAGAGAT GAAGAAGGTC AAGATGTCTT GTTATTCGTT      700
   GATAACATTT TCAGATTCAC CCAAGCCGGT TCAGAAGTTT CTGCCTTATT      750
25 AGGTCGTATT CCATCTGCTG TCGGTATCA ACCAACTTTA GCAACTGATA      800
   TGGGTTTGTT ACAAGAACGT ATTACCACCA CACAAAAGG TTCAGTTACT      850
   TCTGTCCAAG CTGTTTATGT CCCAGCTGAT GATTTAACAG ATCCTGCTCC      900
   AGCTACCACT TTCGCCCATT TGGATGCTAC TACTGTCTTG TCTCGTGGTA      950
   TTTCAGAATT AGGTATTTAC CCAGCTGTCG ATCCATTAGA TTCTAAATCA      1000
30 AGATTATTAG ATGCTTCAGT TGTTGGTCAA GAACATTATG ATGTTGCTAC      1050
   CAACGTTCAA CAAACTTTAC AAGCTTACAA ATCTTTACAA GATATTATTG      1100
   CTATTTTAGG TATGGATGAA TTGTCTGAAC AAGATAAATT GACTGTGCGA      1150
   AGAGCAAGAA AAATCCAAAG ATTCTTATCT CAACCATTG CTGTTGCCGA      1200
   AGTTTTCACT GGTATCCCAG GTAGATTGGT TAGATTAAAA GACACTATCA      1250
35 AATCATTCOA AGATGTTTTG GAAGGTAAAT ATGATCACTT A              1291

```

2) INFORMATION FOR SEQ ID NO: 490

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 508 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 45 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Rhodotorula minuta*  
 (B) STRAIN: ATCC 10658

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 490

```

55 CGTATTGATT CGTGAGTGGC CGTTCCTTA CAGCAAGCTT ATAAAGGAGC      50
   GAAAAAGATC TGACATTCGG CTTATGTGCT ATACAGAGGA ACTCATCAAC      100
   AACGTCGCCA AGGCTCATGG TGGTTACTCT GTCTTCACCG GTGTCGGAGA      150
   GCGAACACGT GAAGGTAACG ATCTCTACCA CGAAATGATT GAAACCGGTG      200
   TCATTCACT CAAGAACGAC AAGTCCAAGG CCGCTCTGGT CTTGCGACAG      250
60 ATGAACGAGC CCCCCGAGC TCGTGCCCGT GTCGCTCTGA CCGGTCTCAC      300

```

CATCGCCGAG TACTTCCGTG ACGTCGAAGG ACAGGATGTG CTACTCTTCA 350  
 TCGACAACAT TTTCCGATTC ACCCAGGCCG GTTCAGAGGT ATCTGCCCTG 400  
 CTCGGACGTA TCCCATCTGC TGTCGGATAC CAGCCCACAC TCTCAACCGA 450  
 TATGGGTGGT ATGCAAGAGC GAATCACAAC CACCAAGAAG GGTTCGATTA 500  
 5 CCTCCGTC 508

## 2) INFORMATION FOR SEQ ID NO: 491

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 686 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 15 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM: *Rhodotorula mucilaginosa*  
 (B) STRAIN: ATCC 66034

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 491

25 TGTCTCATT CAGGAGCTCA TCAACAACAT CGCCAAGGCC CACGGTGGTT 50  
 ACTCGGTCTT CACCGGTGTC GGCGAGCGTA CCCGTGAGGG TAACGGTGAG 100  
 TCTCCCCCTT CAAACTTTTG GCCGGCTAGT TGGCGCAGCG CAAACTGACG 150  
 CGCGCGCCCT GTCCAGACTT GTACCACGAG ATGATCGAGA CTGGTGTCAT 200  
 CCAGCTCGAG AACGACAAC CGAAGTGCGC TCTCGTGTTT GGCCAGATGA 250  
 30 ACGAGCCCCC TGGTGCCCGT GCGCGTGTCG CTCTCACTGG GTTCGTCTT 300  
 TCTCTCTCTC GAGCGTCCTG GCTTGATACG GAACGCTGAC ACGTCACGCA 350  
 GTCTCACTAT TGCTGAGTAC TTCCGTGACG AGGAGGGCCA GGACGTGCTC 400  
 CTCTTCATCG ACAACATCTT CCGTTTCACC CAGGGTGAGC CGCTCCGCG 450  
 GGCATTCTCC CGTTTCTTTC GCGCTGACGT CTGTCCCGTA TAGCCGGTTC 500  
 35 GGAGGTGTCT GCCCTTCTCG GACGTATCCC GTCCGCTGTC GGATACCAGC 550  
 CGACTCTCTC GACCGACATG GGTCAGATGC AGGAGCGTAT CACGTAAGTT 600  
 TGGCCGCAGC TCCGTCCGCG GCGCCCTTTG TGTCTGACCG TGTTCACCG 650  
 CTCAGCACCA CCAAGAAGGG CTCGATCACC TGTGTC 686

## 2) INFORMATION FOR SEQ ID NO: 492

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 625 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Sporobolomyces salmonicolor*  
 (B) STRAIN: ATCC 32311

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 492

TGTCTTGATT CAGGAGCTCA TCAACAACGT CGCCAAGGCC CACGGTGGTT 50  
 ACTCCGTTTT CACCGGTGTC GGTGAGCGTA CCCGTGAGGG TAACGGTGAG 100  
 60 CAACTTTTCG CCGACCTGGC CATTCGGAA CGTGCGGACT GACGAGGACC 150

	ACCTCGAACA	GATCTCTACC	ACGAGATGAT	CGAGACCGGT	GTCATTCAGC	200
	TCGACAACGA	CAAGTCGAAG	ACTGCTCTCG	TCTTCGGCCA	GATGAACGAG	250
	CCCCCTGGCG	CCCGTGCCCC	TGTCGCTCTT	ACTGGTCTCA	CCATCGCGGA	300
	GTACTTCCGT	GACGACGAGG	GCCAGGACGT	GCTTCTCTTC	ATCGACAACA	350
5	TCTTCCGTTT	CACCCAGGGT	ACGTTCGATC	CCGCCCCGTCC	AACACGAATG	400
	TCGTGGTGAC	TGACAACCTG	TTGCGCGTGC	AGCCGGTTCG	GAGGTGTCTG	450
	CCCTTCTCGG	TCGTATCCCC	TCCGCTGTCT	GATACCAGCC	CACTCTCTCG	500
	ACCGACATGG	GTGGCATGCA	GGAGCGTATC	ACGTACGCCC	TCTTCTGCTT	550
	TCTCTCGTTT	CGCTCTGCAT	CGTTCACGCA	TGTTTCGCCC	ACAGCACCAC	600
10	CAAGAAGGGT	TCGATCACCT	GTGTC			625

## 2) INFORMATION FOR SEQ ID NO: 493

15

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1211 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

20

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Sporothrix schenckii*  
 (B) STRAIN: WSA-148

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 493

30	TCTCGTCCGT	GGTGCTAAGG	CCACTGACAC	TGGTTCCCCC	ATTACCATCC	50
	CCGTGCGCCC	CGGTACCCCTC	GGTCGCATCA	TGAACGTCAC	CGGTGACCCG	100
	ATCGACGAGC	GCGGTCCCCT	CAAGACCGAC	AAGTTCCGTC	CCATCCACGC	150
	TGAGGCTCCC	GAGTTCGTTG	ACCAGTCGAC	CACCGCTGAG	GTTCTCGTGA	200
35	CTGGTATCAA	GGTCGTCGAT	CTGCTTGCTC	CCTACGCCCC	TGGTGGTAAG	250
	ATTGGTCTGT	TTGGCGGTGC	CGGTGTTGGC	AAGACCGTGT	TCATCCAGGA	300
	GCTCATCAAC	AACATCGCCA	AGGCCACCGG	TGGTTACTCC	GTCTTCACCG	350
	GTGTGCGCGA	GCGTACCCGT	GAGGGTAACG	ATCTGTACCA	CGAAATGCAG	400
	GAGACCTCTG	TCATTACAGT	TGACGGTGAC	TCCAAGGTCG	CCCTGGTGTT	450
40	CGGTGAGATG	AACGAGCCCC	CTGGTGCTCG	TGCCCCGTGC	GCCTTGACCG	500
	GTTTGACTGT	CGCTGAGTAC	TTCGCTGACG	AGGAGGGCCA	GGATGGTATG	550
	TTTTGAATTA	TTTCCTTGTC	GTACAGTTCC	AAATCGAAGA	ATTACTAACT	600
	TGTCAGTGCT	TCTCTTCATC	GACAACATTT	TCCGCTTCAC	CCAGGCCGGT	650
	TCTGAGGTGT	CTGCCCTTCT	GGGTCGTATT	CCCTCCGCTG	TCGGTTACCA	700
45	GCCCCACGCTC	GCCGTGGACA	TGGGTCTGAT	GCAGGAGCGT	ATTACCACCA	750
	CCCGCAAGGG	CTCAATTACC	TCCGTCCAGG	CCGTCTACGT	GCCCCGCTGAC	800
	GATCTGACGG	ATCCCGCCCC	CGCCACCACC	TTCGCCCCATC	TGGACGCCAC	850
	CACTGTGCTG	TCCCCGAGGT	TCTCTGAGCT	GGGTATCTAC	CCCGCTGTCT	900
	ACCCCTCGA	CTCCAAGTCG	CGTATGCTGG	ACCCCGGTAT	TGTCGGTGAC	950
50	GACCACTACG	AGACCGCCAC	TCGCGTCCAG	CAGATCCTCC	AAGAGTACAA	1000
	GTCGCTGCAG	GACATCATCG	CCATTCTGGG	TATGGACGAG	CTGTCTGAGG	1050
	CCGACAAGCT	TACAGTCGAG	CGTGCTCGTA	AGATCCAGCG	TTTCCTGAGC	1100
	CAGCCGTTCA	CGGTGCGGCA	GGTCTTCACT	GGTATCGAAG	GCCAGCTGGT	1150
	CGATCTGAAG	GACACTATCG	CTTCGTTCAA	GGCTATCCTG	AGCGGTGAGG	1200
55	GTGACAGCCT	T				1211

## 2) INFORMATION FOR SEQ ID NO: 494

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1133 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Stephanoascus ciferrii*  
 (B) STRAIN: ATCC 52550

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 494

```

15  TCTTGTTAGA GGCACCCAG TCAAGGACAC TGGTGCTCCA ATTACCATTC      50
    CAGTTGGTAA CGGCACTTTG GGCCGTATCG TCAACGTTCT CGGTGAGCCA      100
    ATTGATGAGC GTGGACCAAGT CAAGGCTGAC AAGTTCAGAC CTATTCACGC      150
    TGAGCCACCA ACCTTCGCTG ACCAGTCCAC CTCTGCCGAG GTTCTTGAGA      200
    CCGGTATTAA GGTTGTCGAC TTGCTTGCCC CTTATGCCAG AGGTGGTAAG      250
20  ATTGGTCTTT TCGGTGGTGC CGGTGTCGGT AAGACTGTGT TCATTCAGGA      300
    GCTTATTAAC AACATTGCTA AGGCCACCG TGGTACTCT GTCTTCACTG      350
    GTGTCGGTGA GCGAACTCGT GAAGGTAACG ATTTGTACCA CGAAATGATG      400
    GAGACCGGTG TCATCAACCT TGAGGGTGAC TCCAAGGTGT CTCTTGTTG      450
    CCGTCAGATG AACGAGCCTC CAGGAGCCCC TGCCCGTGTT GCCTTGACCG      500
25  GTTTGACCAT TGCCGAGTAC TTCAGAGATG AGGAGGGCCA GGATGTCTTG      550
    TTGTTCAATTG ACAACATTTT CCGATTCACC CAGGCCGGTT CTGAGGTCTC      600
    TGCCTTGTTG GGTCGTATCC CATCTGCCGT CGGTACCAA CCAACCTTGG      650
    CTACTGATAT GGGTGGTCTT CAAGAACGTA TTACCACCAC TCAAAGGGT      700
    TCCGTCACCT CTGTCCAGGC TGTCTACGTC CCAGCTGACG ATTTGACTGA      750
30  TCCTGCCCCA GCTACCACCT TCGCCCATTT GGACGCCACC ACCGAATTGT      800
    CCCGATCTAT CTCTGAGTTG GGTATCTACC CAGCTGTCTGA CCCTCTTGGT      850
    TCCAAGTCCC GTCTTTTGGA TGCCTCCGTC GTCGGCCAAG AGCACTACGA      900
    CGTTGCCGCC AACGTCCAAC AGACCTTGCA GGCTTACAAG TCTCTCCAGG      950
    ATATCATTCG CATTTTGGGT ATGGACGAAT TGTCTGAGGC TGATAAGCTC     1000
35  ACTGTCGAGC GTGCTCGTAA GATGCAGAGA TTCCTTTCTC AGCCATTAC      1050
    CGTCGCTGAG GTCTTCACTG GTCTCGAGGG TAGACTCGTT TCTTTGAAGG     1100
    ACACCATCCG ATCCTTCAAG GAGATCCTTG ACG                        1133

```

## 2) INFORMATION FOR SEQ ID NO: 495

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 608 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichophyton mentagrophytes*  
 (B) STRAIN: WSA-225

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 495

```

GGAGTTGATT GTAAGTCATT TGAAACCCAG CCCCAAGAAA CAGAAGCTAG      50
GCGAAAATTG GACAATTGAG CAATTTAGCC ATTGGAGAAA AGAAATTTTCG      100
AGTATTAATT GTTTTATAG AACAACTTG CCAAGGCTCA CGGTGGTTAC      150
60  TCTGTCTTCA CTGGTGTGCG AGAGCGTACC CGTGAAGGAA ACGATCTCTA      200

```



```

CCATGAAATG CAGGAGACCC GTGTCATTCA GCTTGATGGC GAGTCCAAGG 250
TCGCCCTGGT CTTCCGCCAG ATGAACGAGC CCCCAGGTGC CCGTGCCCGT 300
GTTGCTCTTA CTGGTTTGAC CATTGCTGAG TACTTCCGTG ATGAGGAAGG 350
TCAAGACGGT GAGTTTCTTA TGGATAAAAA AAAATTTTTT TTTTTTTTTT 400
5 TTTTTTTCAA GAAATTCATG TTCTAACAAA GTGTATTCTA GTGCTTCTCT 450
TCATCGACAA CATTTTCCGT TTTACTCAGG CTGGTTCCGA AGTGTCTGCC 500
CTGCTTGGTG GTATCCCATC TGCCGTCGGT TACCAACCCA CTCTTGCCGT 550
CGACATGGGT GGTATGCAGG AACGTATTAC CACCACCAAG AAGGGATCCA 600
TTACCTCC 608
10

```

## 2) INFORMATION FOR SEQ ID NO: 496

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 794 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
20
(ii) MOLECULE TYPE: Genomic DNA
(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Wangiella dermatitidis
25    (B) STRAIN: WSA-229

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 496

```

30 GTTTATTCAA GAATCATTG TGGGTGGCAT TCTCATAATG TTTCGGCCAC 50
AATTACTGAT TGAAAATAGA ACAACATTGC AAAGGCTCAT GGTGGTTACT 100
CCGTGTTCAC TGGTGTGCGC GAGCGAACTC GTGAGGGTAA CGACTTGTAC 150
CACGAAATGC AGGAGACCTC TGTCATTAG CTGATGGCG AGTCCAAGGT 200
CGCGCTGGTG TTTGGTCAA TGAACGAACC TCCTGGTGCT CGTGCTCGTG 250
TTGCTCTGAC TGGGTAAGTT GTTCCTTCGC TTCTTGCGCG TATCCACATC 300
35 CCCATCTTGA GAATACGCT GCCACCATGT CATGTGATGT TGGGCTGGTT 350
CTGGTTTTTG GGAGGCCCTC AAGTTCAATT TTTGGATGAC AGCACCAGCT 400
TTACAAGATT ATGCTAACTT AATGGAGTCT TACGGTGGCT GAGTTCTTCA 450
GGGATGAGGA GGGACAGGAT GGTAAGTTTG ATAACAATCT CGTCGGTGTC 500
AATATCGACG GCGTACTCTT CGCATCAAAA AACCAAAGAG GTGGTTTGGT 550
40 GTGAGAAGTG CGCCGGAAT AATGGCAACC ACGTGACAAT GACCACGTGT 600
GGGGCTCCCG TGCTAACACG TGACAGTCTT GCTCTTCATC GACAACATTT 650
TCCGATTAC TCAASCCGGT TCTGARGTGT CTGCCTTGCT TGGTCGTATT 700
CCATCTGCCG TCGGTTACCA ACCCACACTC GCCGTCGACA TGGGTCTCAT 750
45 GCAGGAACGT ATCACCACCA CCGGAAGGG ATCCATCACA TCTG 794

```

## 2) INFORMATION FOR SEQ ID NO: 497

```

50 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1148 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
55
(ii) MOLECULE TYPE: Genomic DNA
(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Yarrowia lipolytica
60    (B) STRAIN: ATCC 38295

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 497

```

5   TCTTGTCCGA GGCACCGCCG TCGCTGACAC CGGTGCTCCC ATCACTATCC      50
    CCGTCGGCCG AGGTACCCTT GGTCGAATCA TCAACGTCTG TGGTGAGCCC      100
    ATTGACGAGC GAGGACCCAT CGAGGCTTCC AAGTACCTCC CCATCCACGC      150
    TGACCCCCCT ACCTTCGCTG AGCAGTCTAC CTCCGCTGAG GTTCTCGAGA      200
    CCGGTATTAA GGTCGTCGAC CTCCTCGCCC CTTACGCCCC AGGTGGTAAG      250
    ATTGGTCTCT TCGGTGGTGC CGGTGTCGGT AAGACTGTCT TCATCCAGGA      300
10  GCTGATTAAAC AACATTGCCA AGGCCCATGG TGGTTTCTCC GTTTTCTGCG      350
    GTGTCGGTGA GCGAACCCGA GAGGGTAACG ATCTTTACCG AGAGATGAAG      400
    GAGACTGGTG TCATCAACCT CGAGGGTGAG TCTAAGGTCA CCCTCGTCTT      450
    GGGTCAGATG AACGAGCCTC CCGGAGCCCC TGCCCCAGTC GCCCTTACTG      500
    GTCTGACCAT TGCCGAGTAC TTCCGAGACG AGGAGGGTCA GGATGTGTTG      550
15  CTCTTCGTTG ACAACATTTT CCGATTCAAC CAGGCCGGTT CCGAGGTGTC      600
    CGCTCTGCTT GGTCGAATTC CCTCCGCTGT CGGTTACCAG CCCACTCTGG      650
    CCACCGATAT GGGTGCCCTC CAGGAGCGAA TTACCACCAC CCAGAAGGGT      700
    TCCGTCACTT CCGTCCAGGC CGTCTACGTG CCTGCCGATG ATTTGACCGA      750
    TCCTGCTCCC GCCACCACCT TCGCCCATCT TGACGCCACC ACCGTCCTGT      800
20  CCCGAGGTAT TTCCGAGCTG GGTATCTACC CCGCTGTCGA TCCCCTTGAT      850
    TCCAAGTCTC GACTTCTGGA TATCGATGTT GTCGGAAGG AGCACTACGA      900
    TGTTGCTTCC AACGTCCAGC AGACCCTCCA GGCTTACAAG TCTCTCCAGG      950
    ATATCATTGC CATTCTTGGT ATGGATGAGC TGTCCGAGCA GGACAAGCTG     1000
    ACCGTCGAGC GAGCTCGAAA GATCCAGCGA TTCCTGTCTC AGCCCTTCAC     1050
25  CGTCGCCGAG GTTTTTCACCG GTATTGAGGG ACGACTTGTC TCTCTCAAGG     1100
    AACTGTCCG ATCCTTCAAG GAGATCCTTG ACGGTAAGCA CGATGCTC      1148

```

## 30 2) INFORMATION FOR SEQ ID NO: 498

## (i) SEQUENCE CHARACTERISTICS:

```

    (A) LENGTH: 966 bases
    (B) TYPE: Nucleic acid
35  (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 40 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Aspergillus fumigatus
    (B) STRAIN: WSA-172

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 498

```

45  GCGCTATTGT CGTTGTTGCT GCCTCCGACG GTCAGATGTA GGTGGAACAT      50
    CTTGGGAAAT ACGTCGTAAA ACACGTCGCT TACGTTTTTCG CGAATAGGCC      100
    CCAGACTCGT GAGCATTTGC TGCTCGCCCC CCAGGTTGGT GTCCAGAAGA      150
    TCGTTGTCTT CGTCAACAAA ATCGATGCTA TTGATGATCC GGAGATGCTG      200
50  GAACTGGTCG AACTCGAGAT GCGTGAGCTG CTGAACAGCT ACGGTTTCGA      250
    GGGTGAAGAG ACTCCGATCA TTTTCGGTTC CGCTCTCTGT GCTCTCGAAG      300
    GACGCCGTGA CGACATCGGT AAAGACAGAA TTGAGCAGCT TATGAACGCT      350
    GTCGACACCT GGATCCCCAC TCCTCAGCGT GACCTCGACA AACCTTTCTT      400
    GATGTCTGTC GAGGAAGTGT TCTCTATCGC CGGCCGTGGT ACCGTGGCTT      450
55  CTGGTCGTGT CGAGCGTGGT ATCTTGAAGA AGGACTCTGA GGTGAGATT      500
    GTTGGAGGCT CCTTCGAACC CAAGAAGACC AAAGTCACCG ACATTGAAAC      550
    CTTCAAGAAG AGCTGTGATG AATCGCGTGC TGGTGACAAC TCTGGTCTCC      600
    TCCTGCGTGG TATCCGACGT GAAGACGTCA AGCGTGGTAT GGTCAATTGCT      650
    GTTCCCGGCA GCACCAAGGC TCACGACAAG TTCCTCGTCT CCATGTACGT      700
60  CCTGACCGAG GCGGAGGGTG GTCGTCGTAC TGGCTTCGGT GCCAACTACC      750

```

GTCCCAAGT CTTTCATCCGT ACTGCAGGTA AGTTCCCGCA CACCGTGTCC 800  
 AGATCTTCCG AGAGATTAGC GATATATGCT AATGATTCAT CAGACGAGGC 850  
 TGCTGACCTC AGCTTCCCTG ACGGCGACCA ATCTCGCAGA GTTATGCCTG 900  
 GTGACAACGT CGAGATGATC CTGAAGACCC ACCACCCTGT TGCTGCTGAG 950  
 5 GCTGGTCAAC GCTTCA 966

2) INFORMATION FOR SEQ ID NO: 499

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 846 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Blastoschizomyces capitatus*  
 (B) STRAIN: ATCC 10663

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 499

25 TGGTGCTATT ATTGTTGTTG CTGCTTCTGA TGGTCAAATG CCCCAAACCC 50  
 GTGAGCACTT GCTTCTTGCT CGTCAAGTTG GTGTTAAGCA CATTGTTGTT 100  
 TTCGTTAACA AGATTGATAC TATTGATGAT CCTGAAATGT TGGAACCTGT 150  
 CGAAATGGAA ATGAGAGAAC TTCTTTCTTC TTACGGTTTT GATGGCGATA 200  
 ACACCCCTGT CATTATGGGT TCTGCTCTCT GTGCTCTTGA AGGTCGTGAA 250  
 30 CCAGAAATG GTGAACAAAG AATCAACCAA CTCCTTGATG CTATCGATGA 300  
 ATACATTCCT ACCCCAGTTC GTGATATGGA CCAACCTTTC TTGATGCCAC 350  
 TTGAAGGTGT TTTCTCTATT CCAGGTCGTG GACTGTTGC CACTGGACGT 400  
 GTCTATCGTG GTACTTTGAA GAGAGGTGAA GAAGTTGAAG TTGTTGGCTA 450  
 CAATGATGCT CCAATCAAGA CCACCGTTAC TGGTATTGAA ATGTTCAAGA 500  
 35 AGGAACCTGA TCAAGCTCAA GCTGGTGACA ACGCTGGTAT TCTTTTGAGA 550  
 GGTGTTAAGC GTGAAGACCT TAAGCGTGGT ATGGTTGTTG CTAAACCAGG 600  
 TACCGTTAAG CCACACACCA AGTTCCTTGC CTCCATCTAT GTTTTGACTA 650  
 AGGAAGAAGG TGGCAGACAC TCTGGCTTTG GTCTTAACCTA CAGACCTCAA 700  
 CTTTTCCTTG GTTCTGCTGA TGTTACCACCT GTCTTGACCT TCCCAGAGGG 750  
 40 TGTTGACCAA AGCACTCAAG TCATGCCAGG TGACAACACT GAAATGGTTT 800  
 GCGAACTTGT TCACCCAGTT GCTGTGGAAC AAGGCCAACG TTTCAA 846

45 2) INFORMATION FOR SEQ ID NO: 500

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 846 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida rugosa*  
 (B) STRAIN: ATCC 96275

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 500

```

GGTGCCATTA TTGTTGTTGC TGCCTCTGAT GGACAGATGC CCCAGACCCG      50
TGAGCATCTT TTGCTTGCCC GCCAAGTCGG TATGCAAAAG GTCGTTGTGT      100
TTGTTAACAA GATTGATACC ATTGATGACC CTGAAATGCT TGATCTTGTC      150
GAGATGGAGA TCGTGAACT GTTGAATGAA TATGACTTCG ATGGAGATAA      200
5  CTCTCCTGTC ATTATGGGCT CTGCTCTTGC TGCTCTTGAG GACAAGAACC      250
CCGAGATTGG TAAGGACCGT ATCATGCAGC TCTTGGACGC TGTTGATGAA      300
TGGATCCCTA CCCCCGAGCG TGACCTTGAC AAGCCTTTCA TGATGCCTAT      350
TGAGGCCTCT TTCTCCATTT CTGGTCGTGG TACTGTTGCC ACTGGCCGTG      400
TCGAGCGTGG TATTCTCAAG AAGGGTGAGG AAGTCGAGAT CGTTGGTTTC      450
10 AACAAAGCAGC CCCTGAAATC TGTTGTTACT GGTATTGAAA TGTTCAAGAA      500
GGAAC TTGAT CAGGCCCAGG GCGGTGATAA TGCTGGTATC TTGCTTCGTG      550
GTATTCGTCG TGAGGACTTG CAGCGTGGTA TGGTTTTGGC CAAGCCTGGA      600
ACTGTTAAGG CTCACACCAA GTTCCTTTCC TCCATCTACG TTCTCTCCAA      650
GGAAGAGGGG GGCCGTCAC TCCCTTTCCG TATGAACTAT CGTCCCCAGA      700
15 TGTTCGTTTC TGCACGTGAT CTCACCGTTA CTCTTACTTT CCCTGAGGGT      750
GTTGAACAGC ACACTCAGGT CTTCCCTGGT GAGAACACCG AGATGGTTGG      800
CGAGCTCGTT CACCCTACTG CTATTGAGGT TGGTCAACGC TTCAAC      846

```

20

## 2) INFORMATION FOR SEQ ID NO: 501

## (i) SEQUENCE CHARACTERISTICS:

```

25 (A) LENGTH: 944 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

30

## (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Coccidioides immitis
    (B) STRAIN: Silveira

```

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 501

```

AGTTGTCGTC GTTGCTGCTT CAGACGGTCA AATGTATGCA ACCGAGAGCA      50
CTCCCGGATC TTGGTTTAAA TGGCACTAAT ATAAGACAGG CCTCAAACCTC      100
GAGAGCATTT ACTTCTCGCC CGTCAGATCG GTATCCAAAA AATCGTCGTC      150
40 TTCGTGAACA AGGTTGATGC CATCGAGGAC AAAGAGATGT TGGAGCTTGT      200
TGAATTGGAG ATGCGTGAAC TCCTAACCAG CTACGGTTTC GAGGGTGAAG      250
AAACTCCCAT CATTTTGGC TCTGCTCTCT GTGCCCTCGA AGGAAGACAA      300
CCCGAGATCG GTGTTACCAA GATTGATGAG CTCTTGACAG CCGTCGACAC      350
CTGGATTCCC ACTCCTCAGC GTGAGACTGA CAAGCCCTTC TTGATGTCCA      400
45 TTGAGGAAGT GTTCTCTATT TCCGGACGAG GAACCGTTGT CTCCGGCCGT      450
GTGGAGCGTG GTATCCTCAA GAAGGACTCC GAAGTTGAAA TTGTCGGCGG      500
TTCGCCCGAG CCAATCAAAA CCAAGGTTAC CGATATCGAA ACCTTTAAGA      550
AGTCTTGCGA CGAGTCTCGC GCTGGTGATA ACTCCGGCTT GCTCCTACGA      600
GGCGTTAAGC GTGAAGATAT TAGCCGTGGC ATGGTCGTCG CTGTACCAGG      650
50 AAGTGTC AAG GCCCATACTG AATTCTTAGT TTCGCTTAC GTCCTCACCG      700
AAGCTGAGGG TGGCGCAAA TCTGGATTCA GCAGCAAGTA CCGCCACAG      750
ATGTTTCATT GCAC TGCCCGG TATGTAATAC TGTGATAATT TCGTTGACAT      800
GGTACTGATT GAATTCTATA GACGAAGCGG CTCAGCTCAG CTGGCCCGGA      850
GAAGATCAAG ACAAGATGGC TATGCCAGGA GACAATATCG AAATGATTTG      900
55 CACCACCTTG CACCAGTTG CCGCCGAGGC TGGCCAGCGA TTCA      944

```

## 60 2) INFORMATION FOR SEQ ID NO: 502

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 849 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Fusarium oxysporum*  
 (B) STRAIN: WSA-212

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 502

```

15  GCTATCATCG TCGTTGCTGC CTCCGATGGA CAGATGCCCC AGACCCGTGA      50
    ACACTTGCTC CTCGCTCGTC AGGTCGGTGT CCAGCGAATT GTCGTCTTCG      100
    TCAACAAGGT CGATGCCATT GATGACCCCG AGATGCTTGA GCTCGTCGAG      150
    ATGGAGATGC GCGAGCTTCT TAACACCTAC GGCTTCGAAG GCGACGACAC      200
    TCCCGTCATC ATGGGCTCTG CTCTTATGTC TCTCCAGAAC CAGCGCCCCG      250
20  AGATTGGCAC CGAGAAGATC GATGAGCTCC TTGCTGCCGT CGACGAGTGG      300
    ATCCCAACCC CCGAGCGTGA CCTTGACAAG CCCTTCCTTA TGTCCGTCGA      350
    GGATGTCTTC TCCATTGCCG GCCGTGGTAC CGTCGTGTCT GGCCGTGTGG      400
    AGCGTGGTAT TCTGAAGCGT GATCAGGAGA TCGAGCTTGT CGGAAAGGGG      450
    CAGGAGGTTA TCAAGACCAA GGTTACCGAC ATTGAGACCT TCAAGAAGTC      500
25  TTGTGAGCAG TCCCAGGCTG GTGACAACTC TGGTCTCCTC ATCCGAGGTG      550
    TTCGCCGTGA GGATGTCCGC CGTGGTATGG TTGTCTGCGC TCCTGGCACC      600
    GTCAAGTCTC ATACCCAGTT TCTCGCTTCC CTCTACGTCC TCACCAAGGA      650
    GGAGGGTGGC CGACACACCG GTTTCAGGA GCACTACCGA CCCCAGCTCT      700
    ATCTCCGAAC TGCAGATGAG TCCATTGACC TGACTTTCCC CGAGGGTACT      750
30  GAGGATGCCT CCAGTAAGAT GGTCAATGCCT GGTGACAACA CCGAGATGGT      800
    TGTCACCATG GGTCAACCCA ATGCCATCGA GGTGAGTCAG CGATTCAAC      849
  
```

## 2) INFORMATION FOR SEQ ID NO: 503

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1064 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Histoplasma capsulatum*  
 (B) STRAIN: G186A5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 503

```

50  TGGTGCTATT GTCGTTGTTG CTGCTGCTGA CGGCCAAATG TAAGACGCCG      50
    CGAGGGACTG CTGAGGGTTT TATGCTTTT AGGCCCCCTT GTTTCTGAGA      100
    GCATGATGAT ACTAATATTC GGAAACGTAT CTATTAGGCC TCAAACACGT      150
    GAGCATTTGC TCCTTGCCCC ACAGGTCGGT GTCCAAAAGA TCGTCGTTTT      200
55  CGTGAACAAA GTCGACGCCC TTGAGGACAA GGAGATGTTG GAGCTTGTCG      250
    AGTTAGAAAT GAGAGAGCTC TTAAACACCT ACGGCTTCGA GGGTGAAGAG      300
    ACACCCATCA TCTTTGGTTC TGCCCTTTGC GCCATGGAAG GCCGTGAGCC      350
    TGAGTTGGGA GAAAAGAAA TTGATGAATT GCTGGAGGCT GTTGATACTT      400
    GGATCCCCAC ACCACAACGT GATACCGAAA AACCTTTCTT GATGTCCGTT      450
60  GAGGAAGTAT TCTCTATCTC CGGCCGTGGA ACCGTTGCCT CCGGTCGTGT      500
  
```

	TGAGCGCGGT	GTCCTCAAGA	AGGATTCAGA	AGTCGAGCTA	ATTGGGGGCG	550
	GCTCCACCCC	CATCAGGACG	AAGGTAAGTG	ATATCGAAAC	TTTCAAGAAA	600
	TCCTGTGACG	AGTCTAGAGC	TGGGGACAAC	TCCGGTCTTT	TATTGCGTGG	650
	TATCAAGCGT	GAAGATATCC	GCCGTGGTAT	GGTAGTTGCC	GTTCCCTGGCA	700
5	GCGTCAAGGC	CCACGACAAG	TTCTTGGTGT	CGATGTATGT	CCTGACCGAA	750
	GCTGAGGGTG	GTCGCCGAAC	CGGATTCGGC	CAGAACTATC	GTCCTCAAAT	800
	GTTTCATCCGC	ACAGCTGGTA	TGTCAAAATG	GGACCCCTTT	TCATAATCCT	850
	TTCTTTTTTT	CCTTTTCCTC	TCTATCTCTC	TTTCTGTTTC	CTTTCAACTC	900
	GCCTGATTCA	CGAAATTAAC	TAACCCGTTT	GATTATAGAC	GAAGCCGCCC	950
10	ATCTCAGCTT	CCCTAGTGGA	GCAGATGAAA	GCAAACTCGT	TATGCCTGGT	1000
	GACAACGTCG	AGATGATCCT	CCAGACACAC	CGCCCCGTGG	CTGCTGAGGC	1050
	CGGCCAGCGA	TTCA				1064

15

## 2) INFORMATION FOR SEQ ID NO: 504

## (i) SEQUENCE CHARACTERISTICS:

	(A)	LENGTH: 982 bases
20	(B)	TYPE: Nucleic acid
	(C)	STRANDEDNESS: Double
	(D)	TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

25

## (vi) ORIGINAL SOURCE:

(A)	ORGANISM: <i>Paracoccidioides brasiliensis</i>
(B)	STRAIN: ATCC 32071

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 504

	TGGTGCTATC	GTCGTTGTTG	CTGCCTCTGA	CGGCCAAATG	TAGGGATTTT	50
	GCAAGACTGG	TGAAAAAATC	TAAAGAAAAT	AGAAAAGATT	GTGCTGATGT	100
	TTGGTATCAG	GCCCCAAACA	CGCGAGCATC	TGCTTCTTGC	CAGACAAGTC	150
35	GGTGTTTCAGA	AAATCGTTGT	TTTCGTCAAC	AAGGTCGATG	CTGTAGAGGA	200
	TAAGGAGATG	TTGGAGCTTG	TCGAATTGGA	GATGAGAGAG	CTCTTGACCA	250
	CCTATGGGTT	CGAGGGTGAG	AAGACACCTA	TCATCTTTGG	TTCTGCGCTC	300
	TGTGCTATGG	AGGGCCGTCA	GCCCCAGTTG	GGAGAGCAGA	AAATTGATGA	350
	ATTACTCGAG	GCTGTGGATA	CTTGATCCC	TACGCCACAG	CGTGATACTG	400
40	ACAAGCCCTT	CCTGATGTCC	ATTGAGGAGG	TGTTCTCTAT	CTCTGGACGA	450
	GGAACCGTTG	CCTCCGGCCG	CGTTGAGCGT	GGTATCCTCA	AGAAGGACTC	500
	CGAAGTTGAA	ATTATTGGCG	GCGGTGTTCC	CACAATCCTG	ACCAAGGTGA	550
	CTGATATCGA	AACCTTCAAG	AAGTCTTGCG	ACGAGTCCAG	AGCCGGGGAC	600
	AACTCCGGCC	TCTTGTTGCG	CGGTGTCAAG	CGTGAGGATA	TCCGCCGTGG	650
45	TATGGTCGTT	GCAGTTCCCG	GAAGCGTCAA	AGCACATGAC	AGATTCTTGG	700
	TGTCGATGTA	CGTTCTGACC	GAGGCTGAGG	GTGGTCGCCG	CACTGGCTTC	750
	GGTCAGAACT	ATCGTCCTCA	AATGTTTCATC	CGCACAGCTG	GTACGTTTCAT	800
	TCTTTCACTA	TATTCCTATA	TGCATAGCCC	GATCCTCCCA	TTAACTAATT	850
	GACACAGACG	AGGCTGCTGA	ACTCAGCTGG	CCTGATGGAG	ACGACGAAGC	900
50	CAAAATGGTC	ATGCCCCGGT	ACAATGTTGA	AATGGTCCTG	AAGTCACACC	950
	GCCCCGGGGC	GCTGAGGCTG	GACAGCGATT	CA		982

## 55 2) INFORMATION FOR SEQ ID NO: 505

## (i) SEQUENCE CHARACTERISTICS:

	(A)	LENGTH: 931 bases
	(B)	TYPE: Nucleic acid
60	(C)	STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii)MOLECULE TYPE: Genomic DNA

5 (vi)ORIGINAL SOURCE:

(A) ORGANISM: *Penicillium marneffei*

(B) STRAIN: ATCC 58950

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 505

10 CGCTGTTGTT GTCGTCGCTG CTTCTGATGG TCAAATGTAA CATATCCACG 50  
 AGCTGCCAAT TATGGACACT GCTGATAAGA ATAGGCCCCA AACCCGTGAG 100  
 CACTTGCTCC TCGCCCGTCA GGTTCGGTGT CAAAAGATCG TCGTCTTCGT 150  
 CAACAAGGTT GATGCCGTCG AGGACCCCGA GATGTTGGAA CTTGTCGAAT 200  
 15 TGGAAATGCG TGAACCTCTG ACCACCTACG GTTTCGAGGG TGAAGAGACC 250  
 CCTATCATTT TCGGATCCGC TCTTTGCGCC TTGGAAGGCC GCAAGCCCGA 300  
 GATTGGCGAA CAGAAGATTG ACGAGCTCAT GAACGCCGTT GATACCTGGA 350  
 TCCCCACCCC CCAGCGTGAC CTTGACAAGC CCTTCTTGAT GTCCGTTGAG 400  
 GAAGTTTTCT CCATCTCTGG TCGTGGTACC GTTGCATCTG GTCGTGTTGA 450  
 20 GCGTGGTATT TTGCGCAAGG ATTCTGAGGT TGAGATTATC GGATAACAGA 500  
 AGAACCTTAT CAAGACCAAG GTTACCGACA TTGAGACCTT CAAGAAGTCT 550  
 TGCGATGAAT CTCGTGCTGG TGACAACTCT GGCTTGCTTC TCCGTGGTAT 600  
 CAAGCGTGAG GACATTTCGT GTGGTATGGT TATCGCTGCT CCTGGAACCA 650  
 CCAAGGCTCA TGACAACTTC TTGGTCTCCA TGTATGTCTT GACTGAGGCT 700  
 25 GAAGGTGGTC GTCGTACTGG ATTCGGCGCC AACTACCGTC CTCAAGCTTT 750  
 CATCCGTACT GCCGGTATGT TCCCTTTCAA AGTCAATTAA TGAGCGATTT 800  
 GCTAACGAGT TATAGATGAG GCTGCTACTC TCAGCTTCCC CCGTGACGAT 850  
 CAGTCCAAGC AGGTCATGCC CGGTGACAAC GTTGAGATGA TCTTGAAGAC 900  
 ACACCGTCCC GTTGCCGCCG AAGCTGGTCA G 931

## 2) INFORMATION FOR SEQ ID NO: 506

35 (i)SEQUENCE CHARACTERISTICS:

(A) LENGTH: 846 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

40 (ii)MOLECULE TYPE: Genomic DNA

(vi)ORIGINAL SOURCE:

(A) ORGANISM: *Pichia anomala*

(B) STRAIN: ATCC 18205

(xi)SEQUENCE DESCRIPTION: SEQ ID NO: 506

50 TGGTGCTATT ATTGTTGTTG CTGCTTCTGA TGGTCAAATG CCTCAAACCA 50  
 GAGAACATTT RTTATTGGCT AGACAAGTTG GTGTTCAACA CATTGTTGTC 100  
 TTTGTTAACA AAGTTGATAC TATTGATGAC CCAGAAATGT TGGAATTAGT 150  
 TGAAATGGAA ATGAGAGAAT TGTTAAGTAC TTATGGTTTT GATGGTGATA 200  
 ACGTCCAGT TGTTATGGGT TCTGCTTTAT GTGCCTTGGA AGGTCGTGAA 250  
 GAAGAAATTG GTGTCAAAGC TATTGATAAA TTATTAGCTG CTGTTGATGA 300  
 55 ATATATCCCA ACCCCACAAA GAGATTTAGA AAAACCATTC TTGATGGGTG 350  
 TTGAAGATGT CTTYTCAATC TCAGGTAGAG GTACCGTTGT TACTGGTCGT 400  
 GTTGAACGTG GTAACCTGAA GAAAGGTGAT GAAGTTGAAA TTGTTGGTTT 450  
 AAACAAACT CCATTGAAAA CTACTGTYAC NGGTATTGAA ATGTTCAAAA 500  
 AAGAATTGGA CCAAGCTATG GCTGGTGATA ACTGTGGTAT CTTATTACGT 550  
 60 GGTATCAAAA GAGATGACAT YAAAAGAGGT ATGGTTATTG CTAAAACCGG 600

	TACCATCTCW	GCTCACACTA	AATTCTTAGC	CTCAATGTAT	ATTTTGACTA	650
	AAGAAGAAGG	TGGTCGTCAC	TCAGGTTTTG	GTGAACATTA	CAGACCTCAA	700
	TTATTCATCA	GAAGTGGTGA	TGTTACCGTT	GTTTTAACCT	TYCCAGAAGG	750
	TGGTGATTCA	TCTCAACAAA	TCTTACCAGG	TGACAATGTC	GAAATGGTTT	800
5	GTGAATTGGT	TCACCCAAC	GCTTTAGAAG	CTGGTCAAAG	ATTCAA	846

## 2) INFORMATION FOR SEQ ID NO: 507

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 964 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichophyton mentagrophytes*
- (B) STRAIN: WSA-225

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 507

25	GGAGCTGTTG	TCGTTGTCGC	AGCTTCTGAC	GGTCAAATGT	AATTGAATGC	50
	CCGCCCAGAC	GGATGAAAGG	ATTTGACGTT	TCTAACATCA	GTCTAGGCCT	100
	CAGACCAGAG	AACATTTGCT	CCTTGCCCGC	CAGGTCGGTG	TCCAGAAGCT	150
	GGTCGTTTTT	GTTAACAAGG	TCGATGCCGT	TGAGGACCCA	GAGATGTTGG	200
	AGCTTGTCGA	ACTTGAAATG	CGTGAACCTC	TCAGCCACTA	CAGTTTTGAG	250
30	GGTGAGGAGA	CCCCATCAT	TTTTGGCTCT	GCTCTCTGTG	CCCTCGAGTC	300
	CCGTCGACCT	GAGCTTGGTG	TCGAGAAGAT	TGACGAGCTA	TTGAACGCCG	350
	TCGACACCTG	GATCCCCACC	CCCGAGCGCG	CCACTGATAA	GCCTTTCCTC	400
	ATGTCCATTG	AGGAAGTGTT	CTCTATCTCT	GGTCGTGGTA	CCGTCGTCTC	450
	CGGTCGTGTT	GAGCGTGGTA	TCCTCAAGAA	GGATTCCGAC	GTCGAAATTG	500
35	TTGGTGCGTC	TACCACCCCT	ATCAAGACCA	AGGTCACAGA	TATCGAAACC	550
	TTCAAGAAGT	CCTGCGATGA	ATCTCGAGCT	GGTGACAAC	CTGGTCTCCT	600
	TCTCCGAGGT	ATCAAGCGTG	AGGACTTGAA	GCGTGGAATG	GTTGTTGCTG	650
	CCCCCGGATC	CACCAAGGCT	CACACCGACT	TCATGGTCTC	CCTCTACGTC	700
	CTGACTGAGG	CTGAGGGTGG	TCGTTCCAAC	GGCTTCACCC	ACAAGTACCG	750
40	CCCTCAAATG	TTCATCCGTA	CTGCTGGTAT	GTAACCAAAG	TTCCCGCTAT	800
	TTACTAAGTA	GATCATTGCT	AACTTGTTAT	CCCTTCCGTA	GACGAAGCCG	850
	CATCTTTCAG	CTGGCCTGGA	GAAGACCAAG	ACAAGAAGGC	TATGCCTGGT	900
	GACAACGTCG	AGATGATTG	CAAAACCCTC	CACCCCATTG	CTGCCGAGGC	950
	TGGCCAACGA	TTCA				964

## 2) INFORMATION FOR SEQ ID NO: 508

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 844 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Yarrowia lipolytica*
- (B) STRAIN: ATCC 38295



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 508

	GGTGCTATCA	TTGTCGTTGC	TGCTGGAGAC	GGTTCCAAGC	CCCAGACCCG	50
5	AGAGCATCTG	CTGCTTGCTC	GACAGGTCGG	TGTCCAGAAC	CTGGTTGTGT	100
	TTGTTAACAA	GGTTGATCAG	ATTGATGATA	AGGAGATTCT	TGAGCTCGTT	150
	GACATGGAGA	TGCGAGATCT	GCTGACCCAG	TACGGTTTTG	ATGGTGACAA	200
	CACCCCCGTT	GTCATGGGCT	CTGCTCTGTG	CGCTCTTGAG	GGCAAGCAGA	250
	AGGATATTGG	AGAGGACGCC	ATCATGGCCC	TTATGGATGC	CGTTGATGAG	300
10	CACATCCCTA	CCCCTAACCG	TGACCTTGAG	AAGCCCTTCC	TGATGCCCCG	350
	TGAGGACGTT	TTCTCCATCT	CTGGCCGAGG	AACTGTTGTT	ACTGGCCGAG	400
	TCGAGCGAGG	AAACCTGAAG	AAGGGTGAGG	AAATCGAGAT	TGTTGGCTAC	450
	AACAACAAGC	CCATCAAGGC	TGTTTGTTACC	GGTATTGAGA	TGTTCAAGAA	500
	GGAGCTCGAG	TCCGCCATGG	CCGGTGACAA	CGCCGGTATC	CTGCTCCGAG	550
15	GTATCAAGCG	AGACGAGATC	AAGCGAGGTA	TGGTCATGTG	CAAGCCTGGC	600
	ACCGTCAACG	CCCACAGCAA	GTTCCCTTGCT	TCTCTTTACA	TCATCCCCAC	650
	CGAGGAGGGT	GGTCGAACCA	GCTCTTTTCGG	CGCCAACACT	CGACCCAGAG	700
	TGTTTCATCCG	AACTTCTTCC	GTCACCGCCA	CTCTCACCTT	CCCCGAGGGT	750
	ACCGACGAGT	CCCAGACCGT	CAACCCCGGT	GACAACACTG	AGATGGTTCT	800
20	CGAGCTTGTT	CACCCTACCG	CCATTGAGGT	CAACCAGCGA	TTCA	844

## 2) INFORMATION FOR SEQ ID NO: 509

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1067 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

30

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

35

- (A) ORGANISM: *Babesia bigemina*  
 (B) STRAIN: Suarez-2

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 509

40	CTTGGACAAG	CTGAAGAGCG	AGCGTGAGCG	TGGTATCACC	ATTGACATTA	50
	CCCTGTGGAA	GTTCGAAACT	GGCAAGTACT	ACTACACCGT	CATTGACGCC	100
	CCCGGTCACC	GTGACTTCAT	TAAGAACATG	ATTACGGGTA	CCTCCCAGGC	150
	CGATGTTGCT	ATGCTTGTCG	TGCCCCGCCG	GGCTGGTGTT	TTCGAAGCTG	200
	CCTTCTCTAA	GGAAGGTCAG	ACCCGTGAGC	ACGCTCTTTT	GGCCTTCACC	250
45	CTTGGTGTC	AGCAGATCAT	TTGCGCCATC	AACAAGATGG	ACAAGTGCGA	300
	CTACAAGGAG	GACCGTTACA	GCGAAATCCA	GAAGGAAGTT	CAGGGTTACC	350
	TGAAGAAGGT	CGGTTACAAC	ATCGAGAAGG	TGCCTTTCGT	CGCCATCTCC	400
	GGTTTCATGG	GTGACAACAT	GGTTGAGCGC	TCCACCAACA	TGCCGTGGTA	450
	CAAGGGCAAG	ACCTTGGTCG	AGGCCCTCGA	CATGATGGAG	CCCCGAAGA	500
50	GGCCCCGTCG	CAAGCCCCTG	CGTCTTCCCC	TCCAGGGTGT	GTACAAGATC	550
	GGTGGTATCG	GTACCGTCCC	TGTCGGTTCG	GTGGAGACTG	GTCAGCTCAA	600
	GGCCGGTATG	GTCTTCACCT	TCGCCCCCAA	CCCGATCACT	ACTGAGTGCA	650
	AATCCGTCGA	AATGCACCAC	GAAGTTATCG	ATGTTGCCAG	CCCTGGTGAC	700
	AACGTTGGTT	TCAACGTGAA	GAACGTGTCC	ACCTCTGACA	TCCGCACTGG	750
55	TCACGTGCTG	TCTGACTCCA	AGAACGACCC	CGCCAAGGCC	GCCGTGTCCT	800
	TCACCGCCCA	GGTCACTATC	TTGAACCAAC	CTGGTACCAT	CAAGGCCGGT	850
	TACTCCCCTG	TGGTTGACTG	CCACACTGCC	CACATCTCGT	GCAAATTCGA	900
	CGAGATCACC	AGCCGTATGG	ACAAGCGTAC	CGGTAAGGCC	CTTGAGGAGA	950
	ACCCCAAGAC	CATCAAGAAC	GGCGACGCCG	CTATGGTTCG	CCTGAAGCCG	1000
60	TGCAAGCCCA	TGGTCGTCGA	GGCCTTCACT	GAATACGCTC	CCCTTGGTCG	1050

TTTCGCCGTG CGTGACG

1067

## 5 2) INFORMATION FOR SEQ ID NO: 510

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1049 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia bovis*  
 (B) STRAIN: Suarez-3

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 510

20 GTGAACGTGA ACGTGGTATT ACTATTGATA TTACCTTATG GAAGTTCGAG 50  
 ACCACCAAGT ACTACTACAC CGTCATTGAT GCCCCTGGTC ACCGTGACTT 100  
 CATCAAGAAC ATGATTACGG GTACTTCTCA AGCCGATGTT GCTATGCTTG 150  
 25 TTGTACCAGC TGAGGCTGGT GGTTCGAGG CCGCTTTCTC CAAGGAAGGA 200  
 CAGACCCGTG AGCACGCTCT TTTGGCTTTC ACCCTTGGTG TCAAACAGAT 250  
 CATCTGTGCC ATTAACAAGA TGGACAAGTG CGACTACAAG GAGGACCGTT 300  
 ACAGTGAAAT CCAGAAGGAA GTCCAGGGTT ACCTCAAGAA GGTCGGTTAC 350  
 AATATTGAGA AGGTGCCCTT CGTTGCCATC TCCGGTTTCA TGGGAGACAA 400  
 CATGGTTGAG CGTTCCACCA ACATGCCCTG GTATAAGGGA AAGACATTGG 450  
 30 TCGAGGCCCT TGATCAGATG GAACCCCAA AGAGGCCCGT TGACAAGCCA 500  
 CTTCTCTTTC CCCTCCAGGG TGTCTACAAG ATCGGTGGTA TCGGTACCGT 550  
 CCCCCTCGGT CGTGTGAAA CTGGTATGTT GAAGGCTGGT ATGATTCTAA 600  
 CCTTTGCTCC TAACCCAATC ACCACTGAAT GCAAATCCGT TGAAATGCAC 650  
 CACGAAACCG TTGAGGTTGC TTACCCCGGT GACAACGTCG GTTCAACGT 700  
 35 AAAGAACGTT TCTACTTCTG ACATTGCGAG TGGTCACGTT GCCTCTGATT 750  
 CTAAGAACGA CCCTGCCAAG GCTGCTGTTT CCTTCACTGC CCAGGTCATT 800  
 GTGCTCAACC ACCCTGGTAC CATTAAAGCC GGTACTGCC CCGTCGTCGA 850  
 TTGCCACACC GCTCACATTT CATGTAAATT CGAAGAGATC ACCAGCCGTA 900  
 TGGACAAGCG TACCGGTAAA TCTCTTGAGG AAAACCCCAA GACCATCAAG 950  
 40 AACGGTGACG CTGCCATGGT TGTGCTCAAG CCAATGAAGC CCATGGTTGT 1000  
 CGAATCCTTC ACTGAGTATG CTCCTCTTGG TCGTTTCGCT GTTCGTGAC 1049

## 45 2) INFORMATION FOR SEQ ID NO: 511

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1070 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Crithidia fasciculata*  
 (B) STRAIN: ATCC 11745

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 511

	TGGACAAGCT	GAAGGCGGAG	CGCGAGCGCG	GTATCACGAT	CGATATTGCC	50
	CTGTGGAAGT	TCGAGTCGCC	CAAGTCCGTG	TTCACGATCA	TCGATGCCCC	100
	CGGCCACCGC	GACTTTCATCA	AGAACATGAT	CACCGGCACC	TCCCAGGCCG	150
5	ATGCCGCCAT	TCTGATGATT	GACTCGACCC	AGGGTGGCTT	CGAGGCTGGC	200
	ATCTCCAAGG	ACGGCCAGAC	CCGCGAGCAC	GCCCTGCTTG	CCTTCACGCT	250
	GGGCGTGAAG	CAGATGGTTG	TGTGCTGCAA	CAAGATGGAC	GACAAGACGG	300
	TGCAGTACGC	CCAGGCCCCG	TACGAGGAGA	TCAGCAAGGA	GGTCGGCGCG	350
	TACCTGAAGC	GCGTGGGCTA	CAACCCGGAG	AAGGTGCGCT	TCATCCCGAT	400
	CTCGGGCTGG	CAGGGCGACA	ACATGATCGA	GAAGTCCGAC	AACATGTCGT	450
10	GGTACAAGGG	TCCCACGCTG	CTGGAGGCGC	TCGACCTGCT	GGAGGCCCCC	500
	GTGCGTCCGG	TGGACAAGCC	GCTGCGCCTG	CCCCTGCAGG	ACGTGTACAA	550
	GATCGGCGGT	ATCGGCCTG	TGCCCCGTGG	CCGTGTGGAG	ACCGGCGTGA	600
	TGAAGCCGGG	TCGAGTTCGA	GTGTTTCGCG	CTGCCAACGT	GACGACCGAG	650
	GTGAAGTCGA	TCGAGATCGA	CCACGAGCAG	CTGGCTGAGG	CCGTGCCCCG	700
15	CGACAACGTG	GGCTTCAACG	TGAAGAACGT	GTCCGTGAAG	GATATCCGCC	750
	GTGGTAACGT	GTGCGGCAAC	ACGAAGAGCG	ACCCCCGAA	GGAGGCGGCC	800
	GACTTCACTG	CCCAGGTGAT	CGTGCTGAAC	CACCCCGGCC	AGATCAGCAA	850
	CGGCTACGCG	CCGGTGCTGG	ACTGCCACAC	GAGCCACATC	GCGTGCAAGT	900
	TCGCGGACAT	CGAGTCCAAG	ATCGACCGCC	GCTCTGGCAA	GGAGCTGGAG	950
20	AAGAGCCCGA	AGGCCATCAA	GTCCGGCGAT	GCGGCCATCG	TGAAGATGAT	1000
	CCCGCAGAAG	CCGATGTGCG	TGGAGGTGTT	CAACGACTAC	CCGCCGCTGG	1050
	GCCGCTTCGC	TGTCCGCGAT				1070

25

## 2) INFORMATION FOR SEQ ID NO: 512

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 1052 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

35

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Entamoeba histolytica*  
 (B) STRAIN: HM1-IMMS

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 512

	GCTGAAAGAG	AAAGAGGAAT	TACTATTGAT	ATTTTCATTAT	GGAAATTTCGA	50
	AACATCTAAA	TACTACTTCA	CTATTATTGA	TGCCCCAGGT	CAYAGAGATT	100
	TCATTAAGAA	CATGATTACT	GGAACTTCAC	AAGCTGATGT	TGCCATCCTT	150
45	ATTGTTGCTG	CTGGTACTGG	WGAATTTGAA	GCTGGTATTT	CAAAGAATGG	200
	ACAAACCAGA	GAACAYATTC	TTCTTTTCATA	CACTCTTGGA	GTAAACAAA	250
	TGATTGTTGG	WGTTAACAAG	ATGGATGCTA	TTCAATATAA	ACAAGAAAGA	300
	TATGAAGAAA	TTAAGAAAGA	AATTAGTGCA	TTCCTTAAGA	AGACWGGATA	350
	TAATCCAGAC	AAGATTCCAT	TTGTCCCAAT	TTCAGGATTC	CAAGGAGATA	400
50	ATATGATTGA	ACCATCAACC	AACATGCCAT	GGTACAAAGG	ACCAACATTA	450
	ATTGGAGCAC	TTGATTTCAGT	CACACCACCA	GAAAGACCAG	TTGATAAACC	500
	ACTTAGACTT	CCACTTCAAG	ATGTTTAYAA	GATTTTCAGGT	ATTGGAACCTG	550
	TACCATGTGG	AAGAGTTGAA	ACTGGAGTTC	TTAGACCAGG	AACTATTGTT	600
	CAATTTGCAC	CATCAGGAGT	TTCATCTGAA	TGTAAATCAG	TTGAAATGCA	650
55	TCACACAGCA	CTTGCTCAAG	CTATTCCAGG	TGATAATGTT	GGATTCAATG	700
	TTAGAAAYTT	AACAGTTAAA	GATATTAAGA	GAGGAAATGT	AGCATCAGAT	750
	GCTAAGAATC	AACCAGCTGT	TGGATGTGAA	GATTTCCTG	CTCAAGTCAT	800
	TGTCATGAAC	CATCCAGGAC	AAATTAGAAA	GCGATATACA	CCAGTTCTTG	850
	ATTGCCATAC	ATCACACATT	GCATGTAAAT	TCGAAGAATT	ATTAAGCAAG	900
60	ATTGATAGAA	GAACAGGTAA	ATCCATGGAA	GGAGGAGAAC	CAGAATATAT	950

TAAGAATGGA GATTCAGCAC TTGTTAAGAT TGTTCCTCAACT AAACCACTTT 1000  
 GTGTTGAAGA ATTTGCTAAA TTCCCACCAT TGGGAAGATT TGCTGTTAGA 1050  
 GA 1052

5

## 2) INFORMATION FOR SEQ ID NO: 513

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1082 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM: *Giardia lamblia*  
 (B) STRAIN: Faubert-1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 513

25 GACGAGCGCG AGCGCGGGAT CACGATCAAC ATCGCGCTCT GGAAGTTCGA 50  
 GACGAAGAAG TACATCGTCA CGATCATCGA CGCCCCGGGC CACCGCGACT 100  
 TCATCAAGAA CATGATCACG GGGACGTCCC AGGCCGACGT CGCGATCCTC 150  
 GTCGTGCGCG CGGGCCAGGG CGAGTTTCGAG GCCGGGATCT CGAAGGACGG 200  
 CCAGACGCGC GAGCACGCGA CCCTTGCGAA CACGCTCGGG ATCAAGACGA 250  
 TGATCATCTG CGTCAACAAG ATGGACGACG GCCAGGTCAA GTACTCGAAG 300  
 GAGCGCTACG ACGAGATCAA GGGCGAGATG ATGAAGCAGC TCAAGAACAT 350  
 30 CGGCTGGAAG AAGGCCGAGG AGTTCGACTA CATCCCGACG TCCGGCTGGA 400  
 CCGGGGACAA CATCATGGAG AAGTCCGACA AGATGCCCTG GTACGAGGGC 450  
 CCGTGCCTGA TCGACGCGAT CGACGGGCTC AAGGCCCGCA AGCGCCCGAC 500  
 CGACAAGCCC CTCCGCCTCC CGATCCAGGA CGTCTACAAG ATCTCGGGCG 550  
 TCGGGACCGT CCCC GCGGGC CGCGTCGAGA CGGGCGAGCT CGCGCCCGGG 600  
 35 ATGAAGGTCG TCTTCGCCCC GACGTCCAG GTCTCGGAGG TCAAGTCCGT 650  
 CGAGATGCAC CACGAGGAGC TCAAGAAGGC CGGGCCCGGG GACAACGTCG 700  
 GCTTCAACGT CCGCGGGCTC GCCGTCAAGG ACCTCAAGAA GGGCTACGTC 750  
 GTCGGGGACG TGACGAACGA CCCGCCGTC GGCTGCAAGA GCTTCACCGC 800  
 CCAGGTCATC GTCATGAACC ACCCGAAGAA GATCCAGCCC GGCTACACGC 850  
 40 CCGTCATCGA CTGCCACACC GCGCACATCG CGTGCCAGTT CCAGCTCTTC 900  
 CTCCAGAAGC TCGACAAGCG CACGCTCAAG CCCGAGATGG AGAACCCGCC 950  
 CGACGCAGGC CGCGGCGATT GCATCATCGT CAAGATGGTC CCCCAGAAGC 1000  
 CCCTGTGCTG CGAGACGTTT AACGACTACG CGCCCCTCGG CCGCTTCGCC 1050  
 45 GTCCGCGACA TGCGCCAAAC CGTTGCCGTC GG 1082

## 2) INFORMATION FOR SEQ ID NO: 514

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1098 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 55 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 60 (A) ORGANISM: *Leishmania tropica*  
 (B) STRAIN: ATCC 30816

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 514

	ACGCGTGGGT	GCTCGACAAG	CTGAAGGCGG	AGCGCGAGCG	CGGCATCACG	50
5	ATCGACATTG	CGCTGTGGAA	GTTCGAGTCG	CCCAAGTCCG	TGTTCACGAT	100
	CATCGATGCG	CCCGGCCACC	GCGACTTCAT	CAAGAACATG	ATCACGGGCA	150
	CGTCGCAGGC	GGACGCCGCC	ATCCTGATGA	TCGACTCGAC	GCATGGTGGC	200
	TTCGAGGCTG	GCATCTCGAA	GGACGGCCAG	ACCCGCGAGC	ACGCGCTGCT	250
	TGCCTTCACT	CTTGCGCTGA	AGCAGATGGT	GGTGTGCTGC	AACAAGATGG	300
10	ACGACAAGAC	GGTGACGTAC	GCGCAGTCGC	GCTACGATGA	GATCAGCAAG	350
	GAGGTGGGCG	CGTACCTGAA	GCGCGTGGGC	TACAACCCGG	AGAAGGTGCG	400
	CTTCATCCCG	ATCTCGGGCT	GGCAGGGCGA	CAACATGATC	GAGAAGTCGG	450
	ACAACATGCC	GTGGTACAAG	GGTCCCACGC	TGCTGGACGC	GCTCGACATG	500
	CTGGAGCCGC	CGGTGCGCCC	GGTGGACAAG	CCGCTGCGCC	TGCCCCTGCA	550
15	GGACGTGTAC	AAGATCGGGC	GTATCGGGAC	GGTGCCCGTG	GGGCGCGTGG	600
	AGACCGGCAT	CATGAAGCCG	GGCGACGTGG	TGACGTTTCG	GCCCGCCAAC	650
	GTGACGACTG	AGGTGAAGTC	GATCGAGATG	CACCACGAGC	AGCTGGCGGA	700
	GGCGCAGCCC	GGCGACAACG	TCGGCTTCAA	CGTGAAGAAC	GTGTCGGTGA	750
	AGGACATCCG	CCGTGGTAAC	GTGTGCGGCA	ACTCGAAGAA	CGACCCGCCG	800
20	AAGGAGGCGG	CCGACTTCAC	GGCGCAGGTG	ATCGTGCTGA	ACCACCCCGG	850
	CCAGATCAGC	AACGGCTACG	CGCCGGTGCT	GGACTGCCAC	ACGAGCCACA	900
	TTGCGTGCCG	CTTCGCGGAA	ATCGAGTCCA	AGATCGACCG	CCGCTCCGGC	950
	AAGGACGTGG	AGAAGAACC	CAAGGCGATC	AAGTCTGGCG	ATGCCGCGAT	1000
	CGTGAAGATG	GTGCCGCGAG	AGCCGATGTG	CGTGGAGGTG	TTCAACGACT	1050
25	ACGCGCCGCT	GGGCCGCTTT	GCCGTGCGCG	ACATGCGCCA	AACCGTTG	1098

## 2) INFORMATION FOR SEQ ID NO: 515

30

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1104 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

35

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

40

- (A) ORGANISM: *Leishmania aethiopica*  
 (B) STRAIN: ATCC 50119

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 515

45	TACGCGTGGG	TGCTCGACAA	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	50
	GATCGACATT	GCGCTGTGGA	AGTTCGAGTC	GCCCAAGTCC	GTGTTACGCA	100
	TCATCGATGC	GCCCGGCCAC	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	150
	ACGTCGCAGG	CGGACGCCGC	CATCCTGATG	ATCGACTCGA	CGCATGGTGG	200
	CTTCGAGGCT	GGCATCTCGA	AGGACGGCCA	GACCCGCGAG	CACGCGCTGC	250
50	TTGCCTTCAC	TCTTGCGCTG	AAGCAGATGG	TGGTGTGCTG	CAACAAGATG	300
	GACGACAAGA	CGGTGACGTA	CGCGCAGTCG	CGCTACGAGG	AGATCAGCAA	350
	GGAGGTGGGC	GCGTACCTGA	AGCGCGTGGG	CTACAACCCG	GAGAAGGTGC	400
	GCTTCATCCC	GATCTCGGGC	TGGCAGGGCG	ACAACATGAT	CGAGAAGTCG	450
	GACAACATGC	CGTGGTACAA	GGGTCCCACG	CTGCTGGACG	CGCTCGACAT	500
55	GCTGGAGCCG	CCGGTGCGCC	CGGTGGACAA	GCCGCTGCGC	CTGCCCCCTG	550
	AGGACGTGTA	CAAGATCGGC	GGTATCGGGA	CGGTGCCCCG	GGGCCGCGTG	600
	GAGACCGGCA	TCATGAAGCC	GGGCGACGTG	GTGACGTTTC	CGCCCGCCAA	650
	CGTGACGACT	GAGGTGAAGT	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	700
	AGGCGCAGCC	CGCGCACAA	GTCGGCTTCA	ACGTGAAGAA	CGTGTCCGTG	750
60	AAGGACATCC	GCCGTGGCAA	CGTGTGCGGC	AACTCGAAGA	ACGACCCGCC	800

```

GAAGGAGGCG GCCGACTTCA CGGCGCAGGT GATCGTGCTG AACCACCCCG      850
GCCAGATCAG CAACGGCTAC GCGCCGGTGC TGGACTGCCA CACGAGCCAC      900
ATTGCGTGCC GCTTCGCGGA AATCGAGTCC AAGATCGACC GCCGCTCCGG      950
CAAGGAGCTG GAGAAGAACC CCAAGGCGAT CAAGTCTGGC GATGCCGCGA     1000
5 TCGTGAAGAT GGTGCCCGAG AAGCCGATGT GCGTGGAGGT GTTCAACGAC     1050
TACGCGCCGC TGGGCCGCTT TGCCGTGCGC GACATGCGCC AAACCGTTGC     1100
CGTC                                     1104

```

10

## 2) INFORMATION FOR SEQ ID NO: 516

## (i) SEQUENCE CHARACTERISTICS:

```

15 (A) LENGTH: 1106 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

20

## (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Leishmania tropica
    (B) STRAIN: ATCC 30815

```

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 516

```

TACGCGTGGG TGCTCGACAA GCTGAAGGCG GAGCGCGAGC GCGGCATCAC      50
GATCGACATT GCGCTGTGGA AGTTCGAGTC GCCCAAGTCC GTGTTACGA      100
TCATCGATGC GCCCGGCCAC CGCGACTTCA TCAAGAACAT GATCACGGGC      150
30 ACGTCGCAGG CGGACGCCGC CATCCTGATG ATCGACTCGA CGCATGGTGG      200
CTTCGAGGCT GGCATCTCGA AGGACGGCCA GACCCGCGAG CACGCGCTGC      250
TTGCCTTCAC KCTTGCGGTG AAGCAGATGG TGGTGTGCTG CAACAAGATG      300
GACGACAAGA CGGTGACGTA CGCGCAGTCG CGCTACGATG AGATCAGCAA      350
GGAGGTGGGC GCGTACCTGA AGCGCGTGGG CTACAACCCG GAGAAGGTGC      400
35 GCTTCATCCC GATCTCGGGC TGGCAGGGGC ACAACATGAT CGAGAAGTCG      450
GACAACATGC CGTGGTACAA GGGTCCCACG CTGCTGGACG CGCTCGACAT      500
GCTGGAGCCG CCGGTGCGCC CGGTGGACAA GCCGCTGCGC CTGCCCCCTGC      550
AGGACGTGTA CAAGATCGGC GGTATCGGGA CCGTGCCCGT GGGGCGCGTG      600
GAGACCGGCA TCATGAAGCC GGGCGACGTG GTGACGTTCG CGCCCGCCAA      650
40 CGTGACGACT GAGGTGAAGT CGATCGAGAT GCACCACGAG CAGCTGGCGG      700
AGGCGCAGCC CGGCGACAAC GTCGGCTTCA ACGTGAAGAA CGTGTGCGTG      750
AAGGACATCC GCCGTGGTAA CGTGTGCGGC AACTCGAAGA ACGACCCGCC      800
GAAGGAGGCG GCCGACTTCA CGGCGCAGGT GATCGTGCTG AACCACCCCG      850
GCCAGATCAG CAACGGCTAC GCGCCGGTGC TGGACTGCCA CACGAGCCAC      900
45 ATTGCGTGCC GCTTCGCGGA AATCGAGTCC AAGATCGACC GCCGCTCCGG      950
CAAGGAGCTG GAGAAGAACC CCAAGGCGAT CAAGTCTGGC GATGCCGCGA     1000
TCGTGAAGAT GGTGCCCGAG AAGCCGATGT GCGTGGAGGT GTTCAACGAC     1050
TACGCGCCGC TGGGCCGCTT TGCCGTGCGC GACATGCGCC AAACCGTTGC     1100
CGTCGG                                     1106
50

```

## 2) INFORMATION FOR SEQ ID NO: 517

55

## (i) SEQUENCE CHARACTERISTICS:

```

    (A) LENGTH: 1099 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
60

```

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania donovani*

5 (B) STRAIN: ATCC 50212

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 517

10	TACGCGTGGG	TGCTCGACAA	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	50
	GATCGACATT	GCGCTGTGGA	AGTTCGAGTC	GCCCAAGTCC	GTGTTACGA	100
	TCATCGATGC	GCCCGGCCAC	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	150
	ACGTCGCAGG	CGGACGCCGC	CATCCTGATG	ATCGACTCGA	CGCATGGTGG	200
	CTTCGAGGCT	GGCATCTCGA	AGGACGGCCA	GACCCGCGAG	CACGCGCTGC	250
	TTGCCTTCAC	GCTTGGCGTG	AAGCAGATGG	TGGTGTGCTG	CAACAAGATG	300
15	GACGACAAGA	CCGTGACGTA	CGCGCAGTCG	CGCTACGATG	AGATCAGCAA	350
	GGAGGTGGGC	GCGTACCTGA	AGCGCGTGCG	CTACAACCCG	GAGAAGGTGC	400
	GCTTCATCCC	GATCTCGGGC	TGGCAGGGCG	ACAACATGAT	CGAGAGGTGC	450
	GACAACATGC	CGTGGTACAA	GGGTCCCACG	CTGCTGGACG	CGCTCGACAT	500
	GCTGGAGCCG	CCGGTGCGCC	CGGTGGACAA	GCCGCTGCGC	CTGCCCCTGC	550
20	AGGACGTGTA	CAAGATCGGC	GGTATCGGGA	CTGTGCCCGT	GGGCCGCGTG	600
	GAGACCGGCA	TCATGAAGCC	GGGCGACGTG	GTGACGTTTC	CGCCCGCCAA	650
	CGTGACGACT	GAGGTGAAGT	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	700
	AGGCGCAGCC	CGGCGACAAC	GTCCGGCTTCA	ACGTGAAGAA	CGTGTCGGTG	750
	AAGGACATCC	GCCGTGGCAA	CGTGTGCGGC	AACTCGAAGA	ACGACCCGCC	800
25	GAAGGAGGCG	GCCGACTTCA	CGGCGCAGGT	GATCGTGCTG	AACCACCCCG	850
	GCCAGATCAG	CAACGGCTAC	GCGCCGGTGC	TGGACTGCCA	CACGAGCCAC	900
	ATTGCGTGCC	GCTTCGCGGA	AATCGAGTCC	AAGATCGACC	GCCGCTCCGG	950
	CAAGGAGCTG	GAGAAGAACC	CCAAGGCGAT	CAAGTCTGGC	GATGCCGCGA	1000
	TCGTGAAGAT	GGTGCCGCAG	AAGCCGATGT	GCGTGGAGGT	GTTCAACGAC	1050
30	TACGCGCCCG	TGGGCCGCTT	TGCCGTGCGC	GACATGCGCC	AAACCGTTG	1099

2) INFORMATION FOR SEQ ID NO: 518

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1098 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

40 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

45 (A) ORGANISM: *Leishmania infantum*

(B) STRAIN: MOU

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 518

50	TACGCGTGGG	TGCTCGACAA	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	50
	GATCGACATT	GCGCTGTGGA	AGTTCGAGTC	GCCCAAGTCC	GTGTTACGA	100
	TCATCGATGC	GCCCGGCCAC	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	150
	ACGTCGCAGG	CGGACGCCGC	CATCCTGATG	ATCGACTCGA	CGCATGGTGG	200
	CTTCGAGGCT	GGCATCTCGA	AGGACGGCCA	GACCCGCGAG	CACGCGCTGC	250
55	TTGCCTTCAC	GCTTGGCGTG	AAGCAGATGG	TGGTGTGCTG	CAACAAGATG	300
	GACGACAAGA	CCGTGACGTA	CGCGCAGTCG	CGCTACGATG	AGATCAGCAA	350
	GGAGGTGGGC	GCGTACCTGA	AGCGCGTGCG	CTACAACCCG	GAGAAGGTGC	400
	GCTTCATCCC	GATCTCGGGC	TGGCAGGGCG	ACAACATGAT	CGAGAGGTGC	450
	GACAACATGC	CGTGGTACAA	GGGTCCCACG	CTGCTGGACG	CGCTCGACAT	500
60	GCTGGAGCCG	CCGGTGCGCC	CGGTGGACAA	GCCGCTGCGC	CTGCCCCTGC	550

	AGGACGTGTA	CAAGATCGGC	GGTATCGGGA	CTGTGCCCCGT	GGGCCGCGTG	600
	GAGACCGGCA	TCATGAAGCC	GGGCGACGTG	GTGACGTTCG	CGCCCGCCAA	650
	CGTGACGACT	GAGGTGAAGT	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	700
	AGGCGCAGCC	CGGCGACAAC	GTCGGCTTCA	ACGTGAAGAA	CGTGTCGGTG	750
5	AAGGACATCC	GCCGTGGCAA	CGTGTGCGGC	AACTCGAAGA	ACGACCCGCC	800
	GAAGGAGGCG	GCCGACTTCA	CGGCGCAGGT	GATCGTGCTG	AACCACCCCG	850
	GCCAGATCAG	CAACGGCTAC	GCGCCGGTGC	TGGACTGCCA	CACGAGCCAC	900
	ATTGCGTGCC	GCTTCGCGGA	AATCGAGTCC	AAGATCGACC	GCCGCTCCGG	950
	CAAGGAGCTG	GAGAAGAACC	CCAAGGCGAT	CAAGTCTGGC	GATGCCGCGA	1000
10	TCGTGAAGAT	GGTGCCGCAG	AAGCCGATGT	GCGTGGAGGT	GTTCAACGAC	1050
	TACGCGCCGC	TGGGCCGCTT	TGCCGTGCGC	GACATGCGCC	AAACCGTT	1098

15 2) INFORMATION FOR SEQ ID NO: 519

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1071 bases  
 (B) TYPE: Nucleic acid  
 20 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

25 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania enriettii*  
 (B) STRAIN: ATCC 50120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 519

30	CTCGACAAGC	TGAAGGCGGA	GCGCGAGCGC	GGCATCACGA	TCGACATTGC	50
	GCTGTGGAAG	TTCGAGTCGC	CCAAGTCTGT	GTTACAGATC	ATCGATGCGC	100
	CCGGCCACCG	CGACTTCATC	AAGAACATGA	TCACCGGCAC	GTCGCAGGCC	150
	GACGCCGCCA	TCCTGATGAT	CGACTCGACC	CAGGGCGGCT	TCGAGGCTGG	200
35	CATCTCGAAG	GACGGCCAGA	CCCGCGAGCA	CGCGCTGCTC	GCCTTCACGC	250
	TTGGTGTGAA	GCAGATGGTG	GTGTGCTGCA	ACAAGATGGA	CGACAAGACG	300
	GTGCAGTACT	CGCAGGCGCG	CTACGAGGAG	ATCAGCAAGG	AGGTGGGCGC	350
	GTACCTGAAG	CGCGTCGGCT	ACAACCCGGA	GAAGGTGCGC	TTCATCCCGA	400
	TCTCGGGCTG	GCAGGGCGAC	AACATGATCG	ACAAGTCGGA	CAGCATGCCG	450
40	TGGTACAAGG	GGCCACGCT	GCTGGACGCG	CTCGACATGC	TGGAGGCGCC	500
	GGTGCGCCCG	GTGGACAAGC	CGTGCCGCTT	GCCCCTGCAG	GACGTGTACA	550
	AGATCGGCGG	TATCGGGACG	GTGCCCCTGG	GCCGCGTGGA	GACCGGCATC	600
	ATGAAGCCTG	GCGACGTGGT	GACGTTTGCG	CCCGCCAACG	TGACGACGGA	650
	GGTGAAGTCG	ATCGAGATGC	ACCACGAGCA	GCTGGCGGAG	GCGGTGCCCG	700
45	GCGACAACGT	CGGCTTCAAC	GTGAAGAACG	TGTCGGTGAA	GGACATCCGC	750
	CGTGGTAAAC	TGTGCGGCAA	CTCGAAGAAC	GACCCGCCGA	AGGAGGCGGC	800
	CGACTTCACG	GCGCAGGTGA	TCGTGCTGAA	CCACCCTGGC	CAGATCAGCA	850
	ACGGCTACGC	GCCGGTGCTG	GACTGCCACA	CGAGCCACAT	CGCGTGCCGC	900
	TTCGCGGATA	TCGAGTCCAA	GATCGACCGC	CGTCTTGCA	AGGAGCTGGA	950
50	GAAGAACCCC	AAGGCGATCA	AGTCCGGCGA	TGCGGCCATC	GTGAAGATGG	1000
	TGCCGCAGAA	GCCGATGTGC	GTGGAGGTGT	TCAACGACTA	CCCGCCGCTG	1050
	GGGCGCTTCG	CTGTGCGCGA	C			1071

55

2) INFORMATION FOR SEQ ID NO: 520

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1071 bases  
 60 (B) TYPE: Nucleic acid



(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania gerbilli*  
(B) STRAIN: ATCC 50121

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 520

```

CTCGACAAGC TGAAGGCGGA GCGCGAGCGC GGCATCACGA TCGACATTGC      50
GCTGTGGAAG TTCGAGTCGC CCAAGTCCGT GTTCACGATC ATCGATGCGC      100
CCGGCCACCG CGACTTCATC AAGAACATGA TCACGGGCAC GTCGCAGGCG      150
15 GACGCCGCCA TCCTGATGAT CGACTCGACG CATGGTGGCT TCGAGGCTGG      200
CATCTCGAAG GACGGCCAGA CCCGCGAGCA CGCGCTGCTT GCCTTCACTC      250
TTGGCGTGAA GCAGATGGTG GTGTGCTGCA ACAAGATGGA CGACAAGACG      300
GTGACGTACG CGCAGTCGCG CTACGATGAG ATCAGCAAGG AGGTGGGCGC      350
GTACCTGAAG CGCGTGGGCT ACAACCCGGA GAAGGTGCGC TTCATCCCGA      400
20 TCTCGGGCTG GCAGGGCGAC AACATGATCG AGAAGTCGGA CAACATGCCG      450
TGGTACAAGG GTCCCACGCT GCTGGACGCG CTCGACATGC TGGAGCCGCC      500
GGTGCGCCCG GTGGACAAGC CGCTGCGCCT GCCCCTGCAG GACGTGTACA      550
AGATCGGCCG TATCGGGACG GTGCCCCGTG CCCGCGTGGG GACCGGCATC      600
ATGAAGCCGG GCGACGTGGT GACGTTCCGC CCCGCCAACG TGACGACTGA      650
25 GGTGAAGTCG ATCGAGATGC ACCACGAGCA GCTGGCGGAG GCGCAGCCCG      700
GCGACAACGT CGGCTTCAAC GTGAAGAACG TGTCGGTGAA GGACATCCGC      750
CGTGGTAACG TGTGCGGCAA CTCGAAGAAC GACCCGCCGA AGGAGGCGGC      800
CGACTTCACG GCGCAGGTGA TCGTGCTGAA CCACCCCGGC CAGATCAGCA      850
ACGGCTATGC GCCGGTGCTG GACTGCCACA CGAGCCACAT TCGTGCCGC      900
30 TTCGCGGAAA TCGAGTCCAA GATCGACCGC CGCTCCGGCA AGGAGCTGGA      950
GAAGAACCCC AAGGCGATCA AGTCTGGCGA TGCCGCGATC GTGAAGATGG     1000
TGCCGCAGAA GCCGATGTGC GTGGAGGTGT TCAACGACTA CGCGCCGCTG     1050
GGCCGCTTTG CCGTGCGCGA C                                     1071

```

35

2) INFORMATION FOR SEQ ID NO: 521

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 1114 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

45 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania hertigi*  
(B) STRAIN: ATCC 50125

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 521

```

TCGTTCAAGT ACGCGTGGGT GCTCGACAAG CTGAAGGCGG AGCGCGAGCG      50
CGGTATCACG ATCGACATTG CGCTGTGGAA GTTCGAGTCG CCCAAGTCGG      100
55 TGTTCACGAT CATCGATGCG CCCGGCCACC GCGACTTCAT CAAGAACATG      150
ATCACCGGCA CGTCGCAGGC GGATGCTGCC ATTCTGATGA TCGATTTCGAC      200
GCAGGGTGGC TTCGAGGCTG GCATCTCGAA GGACGGCCAG ACGCGCGAGC      250
ACGCGCTGCT AGCCTTCACG CTGGGCGTGA AGCAGATGGT TGTGTGCTGC      300
AACAAGATGG ACGACAAGAC GGTGCAGTAC GCGCAGGCGC GCTACGAGGA      350
60 GATCAGCAAG GAGGTGGGCG CGTACCTGAA GCGCGTGGGC TACAACCCGG      400

```

AGAAGGTGCG CTTTCATCCCG ATCTCGGGCT GGCAGGGCGA CAACATGATC 450  
 GAGAAGTCCG ACAACATGTC GTGGTACAAG GGTCCCACGC TGCTGGAGGC 500  
 GCTGGACATG CTGGAGGCGC CCGTGCGCCC GGTGGACAAG CCGCTGCGCC 550  
 TGCCCCTGCA GGACGTGTAC AAGATCGGCG GCATTGGCAC GGTGCCGGTG 600  
 5 GGCCGTGTGG AGACCGGCAT CATGAAGCCC GGCACGTGG TGACGTTCGC 650  
 GCCCGCCAAC GTGACGACGG AGGTGAAGTC GATCGAGATG CACCACGAGC 700  
 AGCTGCAGGA GGCTGTGCCC GGCACAAACG TCGGCTTCAA CGTGAAGAAC 750  
 GTGTGCGGTGA AGGACATCCG CCGTGGTAAC GTGTGTGGCA ACTCGAAGAA 800  
 CGACCCGCCG AAGGAGGCGG CTGACTTCAC GGCACAGGTG ATCGTGCTGA 850  
 10 ACCACCCCGG CCAGATCAGC AACGGCTACG CGCCGGTGCT GGAATGCCAC 900  
 ACCAGCCACA TCGCGTGCCG CTTCCGCGAC ATCGAGTCGA AGATCGACCG 950  
 CCGCTCCGGC AAGGAGCTGG AGAAGAACC CAAGTCCATC AAGTCCGGCG 1000  
 ACGCCGCCAT CGTGAAGATG GTGCCGCAGA AGCCGATGTG CGTGGAGGTG 1050  
 TTCAACGACT ACCCGCCGCT GGGCCGCTTT GCGGTGCGCG ACATGCGCCA 1100  
 15 AACCGTTGCC GTCG 1114

## 2) INFORMATION FOR SEQ ID NO: 522

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1106 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania major*  
 (B) STRAIN: ATCC-50122

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 522

35 TACGCGTGGG TGCTCGACAA GCTGAAGGCG GAGCGCGAGC GCGGCATCAC 50  
 GATCGACATT GCGCTGTGGA AGTTCGAGTC GCCCAAGTCC GTGTTACGA 100  
 TCATCGATGC GCCCGGCCAC CGCGACTTCA TCAAGAACAT GATCACGGGC 150  
 ACGTCGAGG CCGACGCCGC CATCTGTATG ATCGACTCGA CGCATGGCGG 200  
 CTTTCAGGCT GGCATCTCGA AGGACGGCCA GACCCGCGAG CACGCGCTGC 250  
 40 TTGCCTTCAC TCTTGCGGTG AAGCAGATGG TGGTGTGCTG CAACAAGATG 300  
 GACGACAAGA CCGTGACGTA CGCGCAGTCT CGCTACGATG AGATCAGCAA 350  
 GGAGGTGGGC GCGTACCTGA AGCGCGTGGG CTACAACCCG GAGAAGGTGC 400  
 GCTTCATCCC GATCTCGGGC TGGCAGGGCG ACAACATGAT CGAGAAGTCG 450  
 GACAACATGC CGTGGTACAA GGGTCCCACG CTGCTGGACG CGCTCGACAT 500  
 45 GCTGGAGCCG CCGGTGCGCC CCGTGGACAA GCCGCTGCGC CTGCCCCCTGC 550  
 AGGACGTGTA CAAGATCGGC GGTATCGGGA CCGTGCCCGT GGGCCGCGTG 600  
 GAGACCGGCA TCATGAAGCC GGGCGACGTG GTGACGTTCTG CGCCCGCCAA 650  
 CGTGACGACT GAGGTGAAGT CGATCGAGAT GCACCACGAG CAGCTGGCGG 700  
 AGGCGCAGCC CGGCGACAAC GTCGGCTTCA ACCTGAAGAA CGTGTGCGTG 750  
 50 AAGGACATCC GCCGTGGTAA CGTGTGCGGC AACTCGAAGA ACGACCCGCC 800  
 GAAGGAGGCG GCCGACTTCA CGGCGCAGGT GATCGTGCTG AACCACCCCG 850  
 GCCAGATCAG CAACGGCTAT GCGCCGGTGC TGGACTGCCA CACGAGCCAC 900  
 ATTGCGTGCC GCTTCGCGGA AATCGAGTCC AAGATCGACC GCCGCTCCGG 950  
 CAAGGAGCTG GAGAAGAACC CCAAGCGCAT CAAGTCTGGC GATGCCGCGA 1000  
 55 TCGTGAAGAT GGTGCCGCAG AAGCCGATGT CCGTGGAGGT GTTCAACGAC 1050  
 TACGCGCCGC TGGGCCGCTT TGCCGTGCGC GACATGCGCC AAACCGTTGC 1100  
 CGTCCG 1106

## 2) INFORMATION FOR SEQ ID NO: 523

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1105 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania amazonensis*  
 (B) STRAIN: ATCC 50131

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 523

	TACGCGTGGG	TGCTCGACAA	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	50
	GATCGACATT	GCGCTGTGGA	AGTTCGAGTC	GCCGAAGTCC	GTGTTACGA	100
	TCATCGATGC	GCCCGGCCAC	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	150
20	ACGTGCGAGG	CGGACGCGGC	CATCCTGATG	ATCGACTCGA	CGCATGGTGG	200
	CTTCGAGGCT	GGCATCTCGA	AGGACGGCCA	GACCCGCGAG	CACGCGCTGC	250
	TTGCCTTCAC	TCTTGCGGTG	AAGCAGATGG	TGGTGTGCTG	CAACAAGATG	300
	GACGACAAGA	CGGTGATGTA	CGCGCAGTCG	CGCTACGATG	AGATCAGCAA	350
	GGAGGTGAGC	GCGTACCTGA	AGCGCGTGGG	CTACAACCCG	GAGAAGGTGC	400
25	GCTTCATCCC	GATCTCGGGG	TGGCAGGGCG	ACAACATGAT	CGACAAGTCG	450
	GACAACATGC	CGTGGTACAA	GGGTCCACG	CTGCTGGACG	CGCTCGACAT	500
	GCTGGAGCCG	CCGGTGCGCC	CGGTGGACAA	GCCGCTGCGC	CTGCCCCCTGC	550
	AGGACGTGTA	CAAGATCGGC	GGTATCGGGA	CGGTGCCCCG	GGGCCGCGTG	600
	GAGACCGGGA	TCATGAAGCC	GGGCGACGTG	GTGACGTTTC	CGCCCCGCAA	650
30	CGTGACGACT	GAGGTGAAGT	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	700
	AGGCGCAGCC	CGGCGACAAC	GTGGGCTTCA	ACGTGAAGAA	CGTGTGCGTG	750
	AAGGACATCC	GCCGTGGTAA	CGTGTGCGGC	AACTCGAAGA	ACGACCCGCC	800
	GAAGGAGGCG	GCCGACTTCA	CGGCGCAGGT	GATCGTGCTG	AACCACCCCG	850
	GCCAGATCAG	CAACGGCTAC	GCGCCGGTGC	TGGACTGCCA	CACGAGCCAC	900
35	ATCGCGTGCC	GCTTCGCGGA	GATCGAGTCC	AAGATCGACC	GCCGCTCCCG	950
	CAAGGAGCTG	GAGAAGAACC	CCAAGGCGAT	CAAGTCTGGC	GACGCCGCGA	1000
	TCGTGAAGAT	GGTGCCGCGA	AAGCCGATGT	GCGTGGAGGT	GTTCAACGAC	1050
	TACGCGCCGC	TGGGCCGCTT	TGCCGTGCGC	GACATGCGCC	AAACCGTTGC	1100
40	CGTCG					1105

## 2) INFORMATION FOR SEQ ID NO: 524

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1098 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania mexicana*  
 (B) STRAIN: ATCC 50156

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 524

	TACGCGTGGG	TGCTCGACAA	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	50
60	GATCGACATT	GCGCTGTGGA	AGTTCGAGTC	GCCGAAGTCC	GTGTTACGA	100

	TCATCGATGC	GCCCGGCCAC	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	150
	ACGTCGCAGG	CGGACGCGGC	CATCCTGATG	ATCGACTCGA	CGCATGGTGG	200
	CTTCGAGGCT	GGCATCTCGA	AGGACGGCCA	GACCCGCGAG	CACGCGCTGC	250
	TTGCCTTCAC	TCTTGCGTG	AAGCAGATGG	TGGTGTGCTG	CAACAAGATG	300
5	GACGACAAGA	CGGTGATGTA	CGCGCAGTCG	CGCTACGATG	AGATCAGCAA	350
	GGAGGTGAGC	GCGTACCTGA	AGCGCGTGGG	CTACAACCCG	GAGAAGGTGC	400
	GCTTCATCCC	GATCTCGGGG	TGGCAGGGCG	ACAACATGAT	CGACAAGTCG	450
	GACAACATGC	CGTGGTACAA	GGGTCCACG	CTGCTGGACG	CGCTCGACAT	500
	GCTGGAGCCG	CCGGTGCGCC	CGGTGGACAA	GCCGCTGCGC	CTGCCCCTGC	550
10	AGGACGTGTA	CAAGATCGGC	GGTATCGGGA	CGGTGCCCCG	GGGCCGCGTG	600
	GAGACCGGGA	TCATGAAGCC	GGGCGACGTG	GTGACGTTTC	CGCCCGCCAA	650
	CGTGACGACT	GAGGTGAAGT	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	700
	AGGCGCAGCC	CGGCGACAAC	GTCGGCTTCA	ACGTGAAGAA	CGTGTCCGTG	750
	AAGGACATCC	GCCGTGGTAA	CGTGTGCGGC	AACTCGAAGA	ACGACCCGCC	800
15	GAAGGAGGCG	GCCGACTTCA	CGGCGCAGGT	GATCGTGCTG	AACCACCCCG	850
	GCCAGATCAG	CAACGGCTAC	GCGCCGGTGC	TGGACTGCCA	CACGAGCCAC	900
	ATCGCGTGCC	GCTTCGCGGA	GATCGAGTCC	AAGATCGACC	GCCGCTCCGG	950
	CAAGGAGCTG	GAGAAGAACC	CCAAGGCGAT	CAAGTCTGGC	GACGCCGCGA	1000
	TCGTGAAGAT	GGTGCCGCAG	AAGCCGATGT	GCGTGGAGGT	GTTCAACGAC	1050
20	TACGCGCCGC	TGGGCCGCTT	TGCCGTGCGC	GACATGCGCC	AAACCGTT	1098

## 2) INFORMATION FOR SEQ ID NO: 525

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1081 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania tarentolae*  
 (B) STRAIN: II WT

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 525

40	GCTGAAGGCG	GAGCGCGAGC	GCGGCATCAC	GATCGACATT	GCGCTGTGGA	50
	AGTTCGAGTC	GCCCAAGTCG	GTGTTACACG	TCATCGATGC	GCCCGGCCAC	100
	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	ACGTCGCAGG	CGGACGCGGC	150
	CATCCTGATG	ATCGACTCGA	CGCACGGTGG	GTTTCGAGGCT	GGCATCTCGA	200
	AGGACGGGCA	GACGCGCGAG	CACGCGCTGC	TTGCCTTCAC	TCTTGCGCTG	250
45	AAGCAGATGG	TTGTGTGCTG	CAACAAGATG	GACGACAAGA	CGGTGATGTA	300
	CGCGCAGTCG	CGCTACGATG	AGATCAGCAA	GGAGGTGGGC	GCGTACCTGA	350
	AGCGCGTGGG	CTACAACCCG	GAGAAGGTGC	GCTTCATCCC	GATCTCGGGC	400
	TGGCAGGGCG	ACAACATGAT	CGAGAAGTCG	GACAACATGC	CGTGGTACAA	450
	GGGTCCACG	CTGCTGGACG	CGCTCGACAT	GCTGGAGGCG	CCGGTGCGCC	500
50	CGGTGGACAA	GCCGCTGCGC	CTGCCCCTGC	AGGACGTGTA	CAAGATCGGC	550
	GGCATCGGCA	CGGTGCCCGT	GGGCCGCGTG	GAGACCGGCA	TCATGAAGCC	600
	GGGCGACGTG	GTGACGTTTC	CGCCCGCGAA	CGTGACGACG	GAGGTGAAGT	650
	CGATCGAGAT	GCACCACGAG	CAGCTGGCGG	AGGCGCAGCC	CGGCGACAAC	700
	GTCGGCTTCA	ACGTGAAGAA	CGTGTGCGTG	AAGGACATCC	GCCGTGGGAA	750
55	CGTGTGCGGT	AACTCGAAGA	ACGACCCGCC	GAAGGAGGCC	GCCGACTTCA	800
	CGGCGCAGGT	GATCGTGCTG	AACCACCCCG	GCCAGATCAG	CAACGGCTAC	850
	GCGCCGGTGC	TGGACTGCCA	CACGAGCCAC	ATCGCGTGCC	GGTTCGCGGA	900
	CATCGAGTCC	AAGATTGACC	GCCGCTCCGG	CAAGGAGCTG	GAGAAGAACC	950
	CCAAGGCGAT	CAAGTCCGGC	GATGCCGCGA	TCGTGAAGAT	GGTGCCGCAG	1000
60	AAGCCGATGT	GCGTGGAGAT	GTTCAACGAC	TACGCGCCGC	TTGGCCGCTT	1050

TGCTGTGCGC GACATGCGCC AAACCGTTGC C

1081

## 5 2) INFORMATION FOR SEQ ID NO: 526

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1102 bases  
 (B) TYPE: Nucleic acid  
 10 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania tropica*  
 (B) STRAIN: ATCC 50129

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 526

20 AAATACGCGT GGGTGCTCGA CAAGCTGAAG GCGGAGCGCG AGCGCGGCAT 50  
 CACGATCGAC ATTGCGCTGT GGAAGTTCGA GTCGCCCAAG TCCGTGTTCA 100  
 CGATCATCGA TGC GCCCGGC CACCGCGACT TCATCAAGAA CATGATCACG 150  
 GGCACGTCGC AGGCCGACGC GCCCATCCTG ATGATCGACT CGACGCATGG 200  
 25 TGGCTTCGAG GCTGGCATCT CGAAGGACGG CCAGACCCGC GAGCACGCGC 250  
 TGCTTGCTT CACKCTTGGC GTGAAGCAGA TGGTGGTGTG CTGCAACAAG 300  
 ATGGACGACA AGACGGTGAC GTACGCGCAG TCGCGCTACG ATGAGATCAG 350  
 CAAGGAGGTG GGCGCGTACC TGAAGCGCGT GGGCTACAAC CCGGAGAAGG 400  
 TGCGCTTCAT CCCGATCTCG GGCTGGCAGG GCGACAACAT GATCGAGAAG 450  
 30 TCGGACAACA TGCCGTGGTA CAAGGGTCCC ACGCTGCTGG ACGCGCTCGA 500  
 CATGCTGGAG CCGCCGCTGC GCCCGGTGGA CAAGCCGCTG CGECTGCCCC 550  
 TGCAGGACGT GTACAAGATC GCGCGGTATCG GGACGGTGCC CGTGGGGCGC 600  
 GTGGAGACCG GCATCATGAA GCCGGGCGAC GTGGTGACGT TCGCGCCCGC 650  
 CAACGTGACG ACTGAGGTGA AGTCGATCGA GATGCACCAC GAGCAGCTGG 700  
 35 CGGAGGCGCA GCCCGGCGAC AACGTCGGCT TCAACGTGAA GAACGTGTCTG 750  
 GTGAAGGACA TCCGCCGTGG TAACGTGTGC GGCAACTCGA AGAACGACCC 800  
 GCCGAAGGAG GCGGCCGACT TCACGGCGCA GGTGATCGTG CTGAACCACC 850  
 CCGGCCAGAT CCGCAACGGC TACGCGCCGG TGCTGGACTG CCACACGAGC 900  
 CACATTGCGT GCCGCTTCGC GGAAATCGAG TCCAAGATCG ACCGCCGCTC 950  
 40 CGGCAAGGAG CTGGAGAAGA ACCCAAGGC GATCAAGTCT GGCGATGCCG 1000  
 CGATCGTGAA GATGGTGCCG CAGAAGCCGA TGTGCGTGGA GGTGTTCAAC 1050  
 GACTACGCGC CGCTGGGCCG CTTTGCCGTG CCGGACATGC GCCAAACCGT 1100  
 TG 1102

45

## 2) INFORMATION FOR SEQ ID NO: 527

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1105 bases  
 (B) TYPE: Nucleic acid  
 50 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 55 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neospora caninum*  
 (B) STRAIN: Suarez-4

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 527

	GGACRAACTT	AAAGCTGAAC	GTGAGCGTGG	TATCACCATT	GATATCTCCC	50
	TGTGGAAATT	TGAGACCAGC	AAGTACTATG	TTACCATCAT	TGATGCCCCA	100
5	GGACACAGAG	ACTTCATCAA	AAACATGATT	ACAGGCACAT	CCCAGGCTGA	150
	CTGTGCTGTC	CTGATTGTTG	CTGCTGGTGT	TGGTGAATTT	GAAGCCGGTA	200
	TCTCCAAGAA	CGGGCAGACC	CGTGAGCATG	CCCTTNTGGC	TTACACCCTG	250
	GGTGTGAAAC	AACTAATTGT	TGGCGTTAAC	AAAAGKGATT	CCACTGAGCC	300
	ACCCTATAGC	CARAAGAGAT	ACGARGAAAT	TGTTAAGGAA	GTCAGCMCCT	350
10	AYNTTAAAAA	AATTGGYTAC	AACCCCGACA	CAGTANCATT	TGKGCCAATT	400
	TNTGGCTGGA	ATGGTGACAA	CATGCTGGAN	CCAAGTGCTA	ATATGCCATG	450
	GTTCAAGGGA	TGGAAGTCM	CCCGTAAGGA	CGGCAATGCC	AGKGAACCM	500
	CCCTGCTTGA	AGCTYTGGAT	TGCATTYTGC	CACCAAYTTG	CCCAACTGAC	550
	AAACCCTTGC	GTTTGCCTYT	CCAGGATGTC	TATAAAATTG	GKGGTATTGG	600
15	TACTGTCCCT	GTGGGTCGTG	TGGAGACTGG	TGTTCTCAAA	CCTGGCATGG	650
	TGGTCACCTT	TGCTCCAGTC	AATGTAACAA	CTGAAGTGAA	GTCTGTAGAA	700
	ATGCACCATG	AAGCATTGAG	TGAAGCCCTT	CCTGGGGACA	ATGTGGGCTT	750
	CAATGTCAAG	AACGTGTCTG	TCAAAGATGT	CCGTCGTGGC	AATGTGGCTG	800
	GTGACAGCAA	AAATGATCCA	CCCATGGAAG	CTGCTGGCTT	CACAGCTCAG	850
20	GTGATTATTT	TGAACCATCC	AGGCCAAATC	AGTGCTGGAT	ATGCACCTGT	900
	GCTGGATTGT	CACACAGCTC	ACATTGCTTG	CAAGTTTGCT	GAGCTGAAGG	950
	AGAAGATTGA	TCGTCTGTTCT	GGGAAAAGC	TGGAAGATGG	CCCTAAATTC	1000
	TTGAAATCTG	GTGACGCTGC	CATCGTTGAT	ATGGTTCTTG	GCAAGCCCAT	1050
	GTGTGTCGAG	AGCTTCTCTG	ATTATCCTCC	CCTGGGCCGT	TTTGCTGTGC	1100
25	GTGAC					1105

## 2) INFORMATION FOR SEQ ID NO: 528

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 935 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichomonas vaginalis*  
 (B) STRAIN: ATCC 30001

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 528

45	GCACATCCCA	GGCTGATGCT	GCTATCCTTG	TCATCGACTC	CACACTCGGT	50
	GGYTTCTGAAG	CCGGTATCGC	TGAACAAGGC	CAGACACGTG	AACACGCTCT	100
	TCTTGCCTTC	ACACTCGGCA	TCAAGCAGGT	CATTGTGCGC	GTCAACAAGA	150
	TGGATGACAA	GACAGTCAAC	TACAACAAGG	CYCGTTTCGA	CGAAATCACA	200
	GCCGAAATGA	CACGCATCCT	TACAGGCATC	GGCTACAAGC	CAGAAATGTT	250
50	CCGCTTCGTC	CCAATCTCCG	GCTGGGCTGG	CGACAACATG	ACAGAGAAGT	300
	CTCCAAACAT	GCCATGGTAC	AATGGCCCAT	ACCTTCTTGA	AGCCCTCGAT	350
	TCCCTTCAGC	CACCAAAGCG	CCCATTCTGAC	AAGCCACTCC	GTCTTCCACT	400
	CCAGGATGTC	TACAAGATCA	ACGGTATCGG	TACAGTTCCA	GTCGGCCGTG	450
	TCGAATCCGG	CACAATGAAG	CCAGGCATGA	TCGTTAACCT	CGCCCCATCC	500
55	ACAGTTACAG	CTGAAGTTAA	GTCCATCGAA	ATGCACCACG	AATCCCTTCC	550
	AGAGGCTCTT	CCAGGTGACA	ACATCGGCTT	CAACGTCAAG	AACGTTTCCA	600
	CAGTCGATGT	CAAGCGTGGC	TACGTCGTTG	GTGATACAAA	GCGTGACCCA	650
	CCAGTCGAAT	GCGCTTCCTT	CACAGCTCAR	ATGATCATCT	CCAACCACCC	700
	AGGCAAGATC	CACGCCGGCT	ACCAGCCAGT	TTTCGACTGC	CACACAGCTC	750
60	ACATCGCCTG	CAAGTTCGAC	AAGCTCATCC	AGCGTATCGA	TCGTGCGCCAC	800

GGCAAGAAGG	CTACAGAGAA	CCCAGAATAC	ATTCAGAAGG	ATGATGCCGC	850
TATCGTCGAG	GTTGTCCCAT	CCAAGCCACT	CGTCGTCGAG	TCCTTCCAGG	900
AGTACCCACC	ACTCGGCCGT	TTCGCCATCC	GTGAT		935

5

## 2) INFORMATION FOR SEQ ID NO: 529

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1065 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma brucei* subsp. *brucei*  
 (B) STRAIN: EATRO795

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 529

AAGCTGAAGG	CTGAGCGCGA	ACGTGGTATC	ACGATCGACA	TTGCACTGTG	50
GAAATTTCGAG	TCACCCAAGT	CTGTCTTCAC	TATTATTGAT	GCTCCTGGGC	100
25 ACCGTGACTT	CATCAAGAAC	ATGATCACCG	GCACATCGCA	AGCCGACGCA	150
GCCATCCTCA	TCATTGCCCTC	TGCGCAGGGT	GAGTTCGAGG	CTGGTATCTC	200
CAAGGATGGA	CAGACCCGCG	AGCACGCGTT	GCTGGCCTTC	ACTTTGGGTG	250
TGAAGCAGAT	GGTTGTGTGC	TGCAACAAGA	TGGACGACAA	GA CTGTGAAC	300
TACGGACAGG	AGCGGTATGA	CGAGATTGTG	AAGGAGGTGT	CTGCTTACAT	350
30 CAAGAAGGTT	GGGTACAACG	TGGAGAAGGT	GCGCTTCGTC	CCCATCTCCG	400
GATGGCAGGG	CGACAACATG	ATTGAGAAAT	CCGAGAAGAT	GCCATGGTAC	450
AAGGGTCCAA	CGCTCCTGGA	GGCACTAGAC	ATGCTGGAGC	CACCAAGTGC	500
TCCGAGCGAG	AAGCCCCCTG	GTCTGCCACT	GCAGGACGTG	TACAAGATCG	550
GTGGTATTGG	CACCGTGCCC	GTTGGTCGTG	TGGAGACCGG	CGTGATGAAG	600
35 CCTGGTGATG	TGGTGACGTT	TGCCCCCGCC	AACGTGACGA	CCGAGGTGAA	650
ATCGATCGAG	ATGCACCACG	AGCAGCTCGC	TGAGGCGACC	CCCGGTGACA	700
ACGTGGGCTT	TAACGTGAAG	AACGTTTCTG	TAAAGGACAT	CCGCCGTGGC	750
AACGTCTGCG	GTAACACCAA	GAACGACCCC	CCAAAGGAGG	CCGCCGACTT	800
CACGGCACAG	GTGATCATCC	TGAACCAACC	CGGACAGATT	GGAAACGGTT	850
40 ATGCGCCCGT	GCTGGACTGC	CACACATCGC	ACATTGCCTG	CAAGTTCGCG	900
GAGATCGAGT	CGAAGATCGA	CCGTCGCTCT	GGCAAGGAGC	TGGAGAAGGC	950
TCCCAAGTCG	ATCAAGTCTG	GCGACGCCGC	GATCGTGCGC	ATGGTGCCGC	1000
AGAAGCCTAT	GTGCGTGGAG	GTCTTCAACG	ACTACGCGCC	ACTCGGCCGC	1050
TTTGCCGTGC	GTGAC				1065

45

## 2) INFORMATION FOR SEQ ID NO: 530

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1297 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

55

(ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Crithidia fasciculata*  
 (B) STRAIN: ATCC 11745

60

530

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 530

	TTCGCGGAGG	GCGTGCCGCC	GGTGCTGACG	GCGCTGGACG	TGACGGAGGA	50
5	CCTCGGCCGC	GACGAGCCGC	TGACGCTGGA	GATTGTGCAG	CACTTGGACG	100
	CGAACACCGG	CCGCTGCATT	GCCATGCAGA	CGACGGATCT	GCTGAAGCTG	150
	AAGTCGAAGG	TTGTGTCGAC	GGGCGGCAAC	ATCTCCGTGC	CGGTTGGCCG	200
	CGAGACGCTG	GGCCGCATCT	TCAACGTGCT	CGGCGACGCG	ATCGACCAGC	250
	GCGGTGTGGT	GGGCGAGAAG	ATGCGCATGC	CGATCCACGC	CGAGGCGCCG	300
10	AAGCTGGCGG	ACCAGGCCGC	GGAGGACGCG	ATTCTGACGA	CCGGCATCAA	350
	GGTGATCGAC	CTGATTCTGC	CGTACTGCAA	GGGTGGCAAG	ATCGGGCTGT	400
	TCGGCGGTGC	TGGTGTGGGC	AAGACTGTGA	TCATCATGGA	GCTGATCAAC	450
	AACGTGGCCA	AGGGCCACGG	TGGTTTCTCC	GTGTTGCGCC	GCGTTGGCGA	500
	GCGCACCCGC	GAGGGCACGG	ATCTGTACCT	GGAGATGATG	CAGTCGAAGG	550
15	TCATTGACCT	GAAGGGCGAG	TCGAAGTGCG	TGCTGGTGTA	CGGCCAGATG	600
	AACGAGCCCC	CGGGTGCGCG	TGCGCGTGTT	GCGCAGTCTG	CGCTGACGAT	650
	GGCGGAGTAC	TTCCGTGACG	TGGAGGGCCA	GAACGTGCTG	CTGTTTCATCG	700
	ACAACATCTT	CCGCTTCACC	CAGGCCAACT	CCGAGGTGTC	CGCCCTGCTG	750
	GGCCGCATTC	CCGCCGCCGT	GGGCTACCAG	CCGACGCTTG	CCGAGGATCT	800
20	TGGTATGCTG	CAGGAGCGCA	TTACGTGCGC	GACGAAGGGC	TCGATTACGT	850
	CTGTGCAGGC	CGTGTACGTG	CCGGCCGATG	ATATCACGGA	TCCGGCGCCG	900
	GCGACGACCT	TCTCGCACCT	GGATGCGACG	ACGGTGCTGG	ACCGCGCGGT	950
	TGCCGAGTCT	GGCATCTACC	CCGCCGTGAA	CCCGCTGGAG	TGCGCGTCGC	1000
	GTATCATGGA	CCCCGATGTG	ATCGACGTGG	ACCACTACAA	CGTTGCGCAG	1050
25	GATATCGTGC	AGATGCTGAC	CAAGTACAAG	GAGCTGCAGG	ATATCATTCG	1100
	CGTGCTGGGT	ATCGACGAGC	TGAGCGAGGA	GGACAAGCTT	GTGGTGGACC	1150
	GCGCTCGCAA	GGTGACGCGC	TTCTGTTCGC	AGCCGTTCCT	GGTGGCCGAG	1200
	GTGTTACCCG	GCATGACGGG	CCACTACGTG	CAGCTGGAGG	ACACAGTGGA	1250
30	GTCGTTCTCT	GGCCTGCTGA	TGGGCTCGTA	CGACCAGATC	CCGGAGA	1297

## 2) INFORMATION FOR SEQ ID NO: 531

35	(i) SEQUENCE CHARACTERISTICS:
	(A) LENGTH: 1298 bases
	(B) TYPE: Nucleic acid
	(C) STRANDEDNESS: Double
	(D) TOPOLOGY: Linear
40	(ii) MOLECULE TYPE: Genomic DNA
	(vi) ORIGINAL SOURCE:
	(A) ORGANISM: <i>Leishmania tropica</i>
45	(B) STRAIN: ATCC 30816

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 531

	CTTCTCGGAG	GGCGTGCCGC	CCGTGCTGAC	GGCGCTGGAT	GTGACGGAGG	50
50	ACCTTGGCCG	CGATGAGCCG	CTGACGCTGG	AGATCGTGCA	GCACTTGGAC	100
	GCGAACACGG	GCCGCTGCAT	TGCGATGCAG	ACGACGGACC	TGCTGAAGCT	150
	GAAGTCGAAG	GTCGTGTGCA	CCGGCGGCAA	CATCTCTGTG	CCGGTGGGCC	200
	GTGAGACGCT	GGGCCGCATC	TTCAAYGTTT	TGGGCGACGC	GATCGACCAG	250
	CGCGGCCCCG	TGGGCGAGAA	GATGCGCATG	GCGATCCACG	CCGAGGCCCC	300
55	GAAGCTGGCG	GATCAGGCCG	CGGAGGACAC	GATCCTGACG	ACCGGCATCA	350
	AGGTGATCGA	CCTGATTCTG	CCCTACTGCA	AGGGTGGCAA	GATCGGCCTG	400
	TTCCGGCGGTG	CCGGTGTGGG	CAAGACTGTG	ATCATCATGG	AGCTGATCAA	450
	CAACGTCGCG	AAGGGCCACG	GCGGTTTCTC	CGTGTTTGCC	GGCGTTGGCG	500
	AGCGCACGCG	CGAGGGCACG	GACCTGTACC	TGGAGATGAT	GCAGTCGAAG	550
60	GTGATTGACC	TGAAGGGCGA	GTCGAAGTGY	GTGCTTGTGT	ATGGGCAGAT	600



	GAACGAGCCC	CCGGGTGCGC	GCGCGCGCGT	TGCGCAGTCT	GCGCTGACGA	650
	TGGCGGAGTA	CTTCCGCGAC	GTGGAGGGCC	AGAACGTGCT	GCTGTTTCATC	700
	GACAACATCT	TCCGCTTTCAC	GCAGGCGAAC	TCCGAGGTGT	CTGCGCTGCT	750
	GGGCCGCATT	CCGGCCGCCG	TGGGCTACCA	GCCGACGCTT	GCGGAGGATC	800
5	TTGGTATGCT	GCAGGAGCGC	ATCACGTCGA	CAACGAAGGG	GTCGATCACG	850
	TCCGTGCAGG	CCGTGTACGT	GCCAGCGGAT	GATATCACGG	ATCCCGCGCC	900
	CGCGACGACG	TTCTCGCACC	TGGACGCGAC	GACTGTGCTG	GACCGCGCGG	950
	TGGCGGAGTC	GGGCATCTAC	CCTGCCGTGA	ACCCGCTGGA	GTGCGCGTCG	1000
	CGTATCATGG	ACCCTGATGT	GATCGATGTG	GACCACTACA	ACGTTGCGCA	1050
10	GGATATCGTG	CAGATGCTGA	CCAAGTACAA	GGAGCTGCAG	GATATCATTG	1100
	CGGTGCTTGG	CATCGACGAG	CTGAGCGAGG	AAGACAAGGT	TGTTGTGGAC	1150
	CGCGCGCGCA	AGGTGACCCG	GTTCTGTGCG	CAGCCGTTCC	AGGTTGCGGA	1200
	GGTGTTCACG	GGCATGACGG	GCCACTACGT	GCAGCTGGTC	GACACGGTGG	1250
	AGTCGTTCTC	TGGCCTGCTG	ATGGGGTTCG	ACGACCAGAT	CCCGGAGA	1298

15

## 2) INFORMATION FOR SEQ ID NO: 532

## 20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1297 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

25

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania aethiopica*  
 (B) STRAIN: ATCC 50119

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 532

	TTCTCGGAGG	GCGTGCCGCC	CGTGCTGACG	GCGCTGGATG	TGACGGAGGA	50
35	CCTTGGCCGC	GATGAGCCGC	TGACGCTGGA	GATCGTGCAG	CACTTGGACG	100
	CGAACACCGG	CCGCTGCATT	GCGATGCAGA	CGACGGACCT	GCTGAAGCTG	150
	AAGTCGAAGG	TTGTGTCGAC	CGGCGGCAAC	ATCTCTGTGC	CGGTGGGCCG	200
	TGAGACGCTG	GGCCGCATCT	TCAACGTTCT	GGGCGACGCG	ATCGACCAGC	250
	GCGGCCCCGT	GGGCGAGAAG	ATGCGCATGG	CGATCCACGC	CGAGGCCCCA	300
40	AAGCTGGCGG	ATCAGGCCGC	GGAGGACACG	ATCCTGACGA	CCGGCATCAA	350
	GGTGATCGAC	CTGATTCTGC	CCTACTGCAA	GGGTGGCAAG	ATCGGCCTGT	400
	TCGGCGGGTG	CGGTGTGGGC	AAGACTGTGA	TCATCATGGA	GCTGATCAAC	450
	AACGTCGCGA	AGGGCCACGG	TGTTTCTCTC	GTGTTTGCCG	GCGTTGGCGA	500
	GCGCACGCGC	GAGGGCACGG	ACCTGTACCT	GGAGATGATG	CAGTCGAAGG	550
45	TGATTGACCT	GAAGGGCGAG	TCGAAGTGCG	TGCTTGTA	CGGGCAGATG	600
	AACGAGCCCC	CGGGTGCGCG	CGCGCGCGTT	GCGCAGTCTG	CGCTGACGAT	650
	GGCGGAGTAC	TTCCGCGACG	TGGAGGGCCA	GAACGTGCTG	CTGTTCATCG	700
	ACAACATCTT	CCGCTTCACG	CAGGCGAACT	CCGAGGTGTC	TGCGCTGCTG	750
	GGCCGCATTC	CAGCCGCCGT	GGGCTACCAG	CCGACGCTTG	CGGAGGATCT	800
50	TGGTATGCTG	CAGGAGCGCA	TCACGTCGAC	AACGAAGGGG	TCGATCACGT	850
	CCGTGCAGGC	CGTGTAACGT	CCAGCGGATG	ATATCACGGA	TCCCGCGCCC	900
	GCGACGACGT	TCTCGCACCT	GGACGCGACG	ACTGTGCTGG	ACCGCGCGGT	950
	GGCGGAGTCG	GGCATCTACC	CTGCCGTGAA	CCCGCTGGAG	TGCGCGTCTGC	1000
	GTATCATGGA	CCCCGACGTG	ATCGATGTGG	ACCACTACAA	CGTTGCGCAG	1050
55	GATATCGTGC	AGATGCTGAC	CAAGTACAAG	GAGCTGCAGG	ATATCATTGC	1100
	GGTGCTTGGC	ATCGACGAGC	TGAGCGAGGA	AGACAAGGTT	GTTGTGGACC	1150
	GCGCGCGCAA	GGTGACCCGG	TTCCTGTGCG	AGCCGTTCCA	GGTTGCGGAG	1200
	GTGTTACACG	GCATGACGGG	CCACTACGTG	CAGCTGGTCG	ACACGGTGGA	1250
	GTCGTTCTCT	GGCCTGCTGA	TGGGGTTCGTA	CGACCAGATC	CCGGAGA	1297

60

## 2) INFORMATION FOR SEQ ID NO: 533

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1298 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania donovani*  
 (B) STRAIN: ATCC 50212

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 533

5	CTTCTCGGAG	GGCGTGCCGC	CCGTACTGAC	GGCGCTGGAT	GTGACGGAGG	50
20	ACCTTGGCCG	CGATGAGCCG	CTGACGCTGG	AGATCGTGCA	GCACTTGGAT	100
	GCGAACACCG	GCCGCTGCAT	TCCGATGCAG	ACGACGGACC	TGCTGAAGCT	150
	GAAGTCGAAG	GTTGTGTCTGA	CCGGCGGCAA	CATCTCTGTG	CCGGTGGGCC	200
	GTGAGACGCT	GGGCCGCGATC	TTCAACGTTT	TGGGCGACGC	GATCGACCAG	250
	CGCGGCCCCG	TGGGCGAGAA	GATGCGCATG	GCGATCCACG	CCGAGGCCCC	300
25	GAAGCTGGCG	GACCAGGCCG	CGGAGGACAC	GATCCTGACG	ACCGGCATCA	350
	AGGTGATCGA	CCTCATTCTG	CCCTACTGCA	AGGGCGGCAA	GATCGGCCCTG	400
	TTCGGCGGTG	CCGGTGTGGG	CAAGACTGTG	ATCATCATGG	AGCTGATCAA	450
	CAACGTCGCG	AAGGGCCACG	GTGGCTTCTC	CGTGTTTGCC	GGCGTTGGCG	500
	AGCGCACGCG	CGAGGGCACG	GACCTATACC	TGGAGATGAT	GCAGTCGAAG	550
30	GTGATTGACC	TGAAGGGCGA	GTGGAAGTGC	GTGCTTGTTG	ACGGGCAGAT	600
	GAACGAGCCC	CCGGGTGCGC	GCGCGCGCGT	TGCGCAGTCT	GCGCTGACGA	650
	TGGCGGAGTA	CTTCCGCGAC	GTGGAGGGCC	AGAACGTGCT	GCTGTTCATC	700
	GACAACATCT	TCCGCTTCAC	GCAGGCGAAC	TCCGAGGTGT	CTGCGCTGCT	750
	GGGCCGCATT	CCGGCCGCCG	TGGGCTACCA	GCCGACGCTT	GCCGAGGATC	800
35	TTGGTATGCT	GCAGGAGCGC	ATCACATCGA	CGACGAAGGG	GTCGATCACG	850
	TCCGTGCAGG	CCGTGTACGT	GCCGGCGGAT	GATATCACGG	ATCCCGCGCC	900
	CGCGACGACG	TTCTCGCACC	TGGACGCGAC	GACTGTGCTG	GACCGCGCGG	950
	TGGCGGAGTC	GGGCATCTAC	CCTGCCGTGA	ACCCGCTGGA	GTGCGCGTCG	1000
	CGTATCATGG	ACCCCGATGT	GATCGATGTG	GACCACTACA	ACGTTGCGCA	1050
40	GGATATCGTG	CAGATGCTGA	CCAAGTACAA	GGAGCTGCAG	GATATCATTG	1100
	CGGTGCTTGG	CATCGACGAG	CTGAGCGAGG	AGGACAAGGT	TGTGGTGGAC	1150
	CGCGCGCGCA	AGGTGACCCG	GTTCTGTCTG	CAGCCGTTCC	AGGTTGCGGA	1200
	GGTGTTTACG	GGCATGACGG	GCCACTACGT	GCAGCTGGCC	GACACGGTGG	1250
45	AGTCGTTCTC	TGGCCTGCTG	ATGGGGTCGT	ACGACCAGAT	CCCGGAGA	1298

## 2) INFORMATION FOR SEQ ID NO: 534

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1298 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania infantum*  
 (B) STRAIN: MOU

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 534

```

5   CTTCTCGGAG GGCCTGCCGC CCGTACTGAC GGCCTGGAT GTGACGGAGG      50
    ACCTTGGCCG CGATGAGCCG CTGACGCTGG AGATCGTGCA GCACTTGGAT      100
    GCGAACACCG GCCGCTGCAT TGCATGACAG ACGACGGACC TGCTGAAGCT      150
    GAAGTCGAAG GTTGTGTCTGA CCGGCGGCAA CATCTCTGTG CCGGTGGGCC      200
    GTGAGACGCT GGGCCGCATC TTCAACGTTT TGGGCGACGC GATCGACCAG      250
    CGCGGCCCCG TGGGCGAGAA GATGCGCATG GCGATCCACG CCGAGGCCCC      300
10  AAAGCTGGCG GACCAGGCCG CGGAGGACAC GATCCTGACG ACCGGCATCA      350
    AGGTGATCGA CCTCATTCTG CCCTACTGCA AGGGCGGCAA GATCGGCCTG      400
    TTCGGCGGTG CCGGTGTGGG CAAGACTGTG ATCATCATGG AGCTGATCAA      450
    CAACGTCGCG AAGGGCCACG GTGGCTTCTC CGTGTGTTGCC GCGCTTGGCG      500
    AGCGCACGCG CGAGGGCACG GACCTATACC TGGAGATGAT GCAGTCGAAG      550
15  GTGATTGACC TGAAGGGCGA GTCGAAGTGC GTGCTTGTGT ACGGGCAGAT      600
    GAACGAGCCC CCGGTGTGCG GCGCGCGCGT TGCAGAGTCT GCGCTGACGA      650
    TGGCGGAGTA CTTCCGCGAC GTGGAGGGCC AGAACGTGCT GCTGTTCATC      700
    GACAACATCT TCCGCTTCAC GCAGGCGAAC TCCGAGGTGT CTGCGCTGCT      750
    GGGCCGCATT CCGGCCGCCG TGGGCTACCA CCGGACGCTT GCCGAGGATC      800
20  TTGGTATGCT GCAGGAGCGC ATCACATCGA CGACGAAGGG GTCGATCACG      850
    TCCGTGCAGG CCGTGTACGT GCCGGCGGAT GATATCACTG ATCCCGCGCC      900
    CGCGACGACG TTCTCGCACC TGGACGCGAC GACTGTGCTG GACCGCGCGG      950
    TGGCGGAGTC GGGCATCTAC CCTGCCGTGA ACCCGCTGGA GTGCGCGTCTG      1000
    CGTATCATGG ACCCCGATGT GATCGATGTG GACCACTACA ACGTTGCGCA      1050
25  GGATATCGTG CAGATGCTGA CCAAGTACAA GGAGCTGCAG GATATCATTG      1100
    CCGTGCTTGG CATCGACGAG CTGAGCGAGG AGGACAAGGT TGTGGTGGAC      1150
    CGCGCGCGCA AGGTGACCCG GTTCCTGTCT CAGCCGTTCC AGGTTGCGGA      1200
    GGTGTTACAG GGCATGACGG GCCACTACGT GCAGCTGGCC GACACGGTGG      1250
    AGTCGTTCTC TGGCCTGCTG ATGGGGTCGT ACGACCAGAT CCCGGAGA      1298
30

```

## 2) INFORMATION FOR SEQ ID NO: 535

```

35  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1301 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
40
    (ii) MOLECULE TYPE: Genomic DNA
    (vi) ORIGINAL SOURCE:
    (A) ORGANISM: Leishmania gerbilli
45  (B) STRAIN: ATCC 50121

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 535

```

50  GCACTTCTCG GAGGGCGTGC CGCCCGTGCT GACGGCGCTG GATGTGACGG      50
    AGGACCTTGG CCGCGATGAG CCGCTGACGC TGGAGATCGT GCAGCACTTG      100
    GACGCGAACA CCGGCCGCTG CATTGCGATG CAGACGACGG ACCTGCTGAA      150
    GCTGAAGTCG AAGGTTGTGT CGACCGGTGG CAACATCTCT GTGCCGGTGG      200
    GCCGTGAGAC GCTGGGCCGC ATCTTCAACG TTCTGGGCGA TGCGATCGAC      250
    CAGCGCGGCC CCGTGGGCGA GAAGATGCGC ATGGCGATCC ACGCCGAGGC      300
55  CCCGAAGCTG CCGGATCAGG CCGGAGGGA CACGATCCTG ACGACCGGCA      350
    TCAAGGTGAT CGACCTGATT CTGCCCTACT GCAAGGGTGG CAAGATCGGY      400
    CTGTTCCGCG GTGCCGGTGT GGGCAAGACT GTGATCATCA TGGAGCTGAT      450
    CAACAACGTC GCGAAGGGCC ACGGTGGTTT CTCCGTGTTT GCCGGCGTTG      500
    GCGAGCGCAC GCGCGAGGGC ACGGACCTGT ACCTGGAGAT GATGCAGTCG      550
60  AAGGTGATTG ACCTGAAGGG CGAGTCGAAG TGCCTGCTTG TGTACGGGCA      600

```

```

GATGAACGAG CCCCCGGGTG CGCGCGCGCG CGTTGCGCAG TCTGCGCTGA 650
CGATGGCGGA GTACTTCCGC GACGTGGAGG GCCAGAACGT GCTGCTGTTC 700
ATCGACAACA TCTTCCGCTT CACGCAGGCG AACTCCGAGG TGTCCGCGCT 750
GCTGGGCCGC ATTCGGGCCG CCGTGGGCTA CCAGCCGACG CTTGCGGAGG 800
5 ATCTTGGTAT GCTGCAGGAG CGCATCACGT CGACAACGAA GGGGTTCGATC 850
ACGTCCGTGC AGGCCGTGTA CGTGCCAGCG GATGATATCA CGGATCCCGC 900
GCCCCGACG ACGTTCTCGC ACCTTGACGC GACGACTGTG CTGGACCGCG 950
CGGTGGCGGA GTCGGGCATC TACCCTGCCG TGAACCCGCT GGAGTGC GCG 1000
TCGCGTATCA TGGACCCCGA TGTGATCGAT GTGGACCACT ACAACGTTGC 1050
10 GCAGGATATC GTGCAGATGC TGACCAAGTA CAAGGAGCTG CAGGACATCA 1100
TTGCGGTGCT TGGCATCGAC GAGCTGAGCG AGGAAGACAA GGTGTGTTG 1150
GACCGCGCGC GCAAGGTGAC CCGGTTCTTG TCGCAGCCGT TCCAGGTTGC 1200
GGAGGTGTTT ACGGGCATGA CGGGCCACTA CGTGCAGCTG GTCGACACGG 1250
TGGAGTCGTT CTCTGGCTTG CTGATGGGGT CGTACGACCA GATCCCGGAG 1300
15 A 1301

```

## 2) INFORMATION FOR SEQ ID NO: 536

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1298 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania hertigi*  
 (B) STRAIN: ATCC 50125

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 536

```

35 CTTGCGGGAG GGCCTGCCGC CCGTGCTGAC GTCGCTGGAT GTGACGGAGA 50
ACCTCGGCCG CGATGAGCCG CTGACGCTGG AGATTGTGCA GCACTTGGAC 100
GCGAACACCG GTCGCTGCAT TGCGATGCAG ACGACGGACC TGCTGAAGCT 150
GAAGTCGAAG GTCGTGTCGA CCGGTGGCAA CATCTCTGTG CCTGTTGGCC 200
GCGAGACGCT GGGTTCGCAT TTCAACGTGC TTGGCGATGC GATTGACCAG 250
40 CGCGGCCCTG TGGGTGAGAA GATGCGCATG GCGATCCACG CCGAGGCGCC 300
GAAGCTGGCG GATCAGGCGG CAGAGGACAC GATCCTGACG ACCGGCATCA 350
AGGTGATCGA TCTTATTCTG CCGTACTGCA AGGGTGGTAA GATCGGTCTG 400
TTGCGTGTTG CCGGTGTAGG CAAGACTGTG ATTATTATGG AGCTGATCAA 450
TAACGTGGCG AAGGGCCACG GTGGGTTTTC CGTGTTTGCT GGCCTGGGCG 500
45 AGCGCACGCG CGAGGGCACT GACCTGTACC TGGAGATGAT GCAGTCGAAG 550
GTGATTGACC TGAAGGGCGA ATCAAAGTGC GTGCTTGTGT ACGGACAGAT 600
GAACGAGCCC CCGGTGCGC GTGCGCGCGT TGCGCAGTCT GCGCTGACGA 650
TGGCCGAGTA CTTCCGCGAT GTGGAGGGCC AGAACGTGCT GCTGTTCATT 700
GACAACATCT TCCGCTTCAC GCAGGCGAAC TCCGAGGTGT CTGCGCTGCT 750
50 GGGTCGCATT CCTGCCGCCG TGGGCTACCA GCCGACGCTT GCGGAGGATC 800
TGGGCATGCT GCAGGAGCGC ATTACGTCGA CGACGAAGGG CTCGATTACG 850
TCTGTGCAGG CCGTGTACGT GCCTGCGGAT GATATCACGG ACCCGGCGCC 900
CGCGACGACG TTCTCGCACC TGGACGCGAC GACTGTGCTG GACCGTGC GG 950
TGGCAGAGTC GGGCATTTAC CCTGCGGTGA ACCCGCTGGA GTGCGCGTCG 1000
55 CGTATCATGG ACCCGGATGT GATCGATGTG GACCACTACA ACGTTGCGCA 1050
GGATATCGTG CAGATGCTGA CCAAGTACAA GGAGCTGCAG GATATCATTTG 1100
CCGTGCTTGG TATCGACGAG CTGACGAGG AGGACAAGGT TGTGGTGGAC 1150
CGCGCGCGCA AGTGACCCG GTTCCTGTCG CAGCCGTTCC AGGTTGCGGA 1200
GGTGTTCACT GGCATGACGG GTCACTACGT TCAGCTGGAG GACACGGTGG 1250
60 AGTCGTTCTC TGGCCTACTG ATGGGGTCAT ACGACCAGAT CCCGGAGA 1298

```

## 2) INFORMATION FOR SEQ ID NO: 537

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1297 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania major*  
 (B) STRAIN: ATCC 50122

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 537

```

20  CTTCTCGGAG GGCCTGCCGC CCGTGCTGAC GGCCTGGAT GTGACGGAGG      50
    ACCTTGGCCG TGATGAGCCG CTGACGCTGG AGATCGTGCA GCACTTGGAC      100
    GCGAACACCG GCCGCTGCAT TCGATGACG ACGACGGACC TGCTGAAGCT      150
    GAAGTCGAAG GTTGTGTCTGA CCGGCGGCAA CATCTCTGTG CCGGTGGGCC      200
    GTGAGACGCT GGGCCGCATC TTCAACGTTT TGGGCGATGC GATCGACCAG      250
25  CGCGGCCCCG TGGGCGAGAA GATGCGCATG GCGATCCACG CCGAGGCCCC      300
    GAAGCTGGCG GATCAGGCCG CAGAGGACAC GATCCTGACG ACCGGCATCA      350
    AGGTGATCGA CTTGATCCTG CCCTACTGCA AGGGTGGCAA GATCGGCCTG      400
    TTCGGCGGTG CCGGTGTGGG CAAGACTGTG ATCATCATGG AGCTGATCAA      450
    CAATGTGCGG AAGGGCCACG GTGGTTTCTC CGTGTGTTGCC GGCGTTGGCG      500
30  AGCGCACGCG CGAGGGCACG GACCTGTACC TGGAGATGAT GCAGTCGAAG      550
    GTGATTGACC TGAAGGGCGA GTCGAAGTGC GTGCTTGTGT ACGGGCAGAT      600
    GAACGAGCCC CCGGGTGCGC GCGCGCGCGT TGCAGTCT GCGCTGACGA      650
    TGGCGGAGTA CTTCCGCGAC GTGGAGGGCC AGAACGTGCT GCTGTTCATC      700
    GACAACATCT TCCGCTTCAC GCAGGCGAAC TCCGAGGTGT CCGCGCTGCT      750
35  GGGCCGCATT CCGGCCGCCG TGGGCTACCA GCCGACGCTT GCGGAGGATC      800
    TTGGTATGCT GCAGGAGCGC ATCACGTCGA CAACGAAGGG GTCGATCACG      850
    TCCGTGCAGG CCGTGTACGT GCCAGCGGAT GATATCACGG ATCCCGCGCC      900
    CGCGACGACG TTCTCGCACC TGGATGCGAC GACTGTGCTG GACCGCGCGG      950
    TGGCGGAGTC GGGCATCTAC CCTGCCGTGA ACCCGCTGGA GTGCGCGTCG     1000
40  CGTATCATGG ACCCCGATGT GATCGATGTG GACCACTACA ACGTTGCGCA     1050
    GGATATCGTG CAGATGCTGA CCAAGTACAA GGAGCTGCAG GACATCATTG     1100
    CCGTGCTTGG CATCGACGAG CTGAGCGAGG AAGACAAGGT TGTGGTGGAC     1150
    CGCGCGCGCA AGGTGACCCG GTTCCTGTCT CAGCCGTTCC AGGTTGCGGA     1200
    GGTGTTTACG GGCATGACGG GCCACTACGT GCAGCTGGTC GACACGGTGG     1250
45  AGTCGTTCTC TGGCCTGCTG ATGGGGTCGT ACGACCAGAT CCCGGAG      1297

```

## 2) INFORMATION FOR SEQ ID NO: 538

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1297 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania amazonensis*

(B) STRAIN: ATCC 50131

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 538

```

5  TTCTCGGAGG GCGTGCCGCC CGTGCTGACG GCGCTGGATG TGACGGAGGA      50
   CCTTGGCCGC GATGAGCCGC TGACGCTGGA GATCGTGCAG CACCTGGACG      100
   CGAACACCGG CCGCTGCATT GCGATGCAGA CGACGGACCT GTTGAAGCTG      150
   AAGTCGAAGG TTGTGTCGAC CGGCGGCAAC ATCTCTGTGC CGGTGGGCCG      200
   TGAGACGCTG GGCCGCATCT TCAACGTGCT GGGCGACGCG ATCGACCAGC      250
10  GCGGCCCCGT GGGTGAGAAG ATGCGCATGG CGATCCACGC CGAGGCCCCG      300
   AAGCTGGCGG ATCAGGCCGC GGAGGACACG ATCCTGACGA CCGGCATCAA      350
   GGTGATCGAC CTGATTCTGC CCTACTGCAA GGGTGGCAAG ATCGGCCTGT      400
   TTGGTGGCGC CGGTGTGGGC AAGACCGTGA TCATCATGGA GTTGATTAAC      450
   AACGTCGCGA AGGGCCACGG TGGTTTCTCG GTGTTTGCCG GCGTTGGCGA      500
15  GCGCACGCGC GAGGGCACGG ACCTGTACCT GGAGATGATG CAGTCGAAGG      550
   TGATTGACCT GAAGGGCAGG TCGAAGTGCG TGCTTGTTGTA CGGGCAGATG      600
   AACGAGCCCC CGGGTGCGCG CGCGCGCGTT GCGCAGTCTG CGCTGACGAT      650
   GCGGGAGTAC TTCCGAGACG TGGAGGGCCA GAATGTGCTG CTGTTTCATCG      700
   ACAACATCTT CCGCTTCACG CAGGCGAACT CCGAGGTGTC TGCCTGCTG      750
20  GGCCGCATTC CGGCCGCCGT GGGCTACCAG CCGACGCTTG CGGAGGATCT      800
   TGGTATGCTG CAGGAGCGCA TCACGTCGAC GACGAAGGGG TCGATCACGT      850
   CCGTGCAGGC CGTGTACGTG CCTGCGGATG ATATCACGGA TCCGGCGCCC      900
   GCGACGACGT TCTCGCACCT GGACGCGACG ACTGTGCTGG ACCGCGCGGT      950
   GCGCGAGTCG GGGATCTACC CTGCCGTGAA CCCGCTGGAG TGCCTGCTGC     1000
25  GTATCATGGA CCCCAGTGTG ATCGACGTGG ACCACTACAA CGTTGCGCAG     1050
   GATATCGTGC AGATGCTGAC CAAGTACAAG GAGCTGCAGG ATATCATTGC     1100
   GGTGCTTGGT ATCGACGAGC TGAGCGAGGA GGACAAGGTC GTGGTGGACC     1150
   GCGCGCGCAA GGTGACCCGG TTCTGTGTCG AGCCGTTCCA GGTTGCGGAG     1200
   GTGTTACGGG GCATGACGGG CCACTACGTG CAGCTGGCCG ACACGGTGGA     1250
30  GTCGTTCTCT GGGCTGCTGA TGGGGTCGTA CGACCAGATC CCGGAGA       1297

```

2) INFORMATION FOR SEQ ID NO: 539

```

35  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 27 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Single
40  (D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 539

```

45  CCITACATCC TBGTYGCICT IAACAAG

```

27

50 2) INFORMATION FOR SEQ ID NO: 540

```

      (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 23 bases
      (B) TYPE: Nucleic acid
55  (C) STRANDEDNESS: Single
      (D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 540

GGDGCITCYT CRTCGWAITC CTG

23

5

## 2) INFORMATION FOR SEQ ID NO: 541

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 541

GTKGAAATGT TCCGCAAGCT GCT

23

20

## 2) INFORMATION FOR SEQ ID NO: 542

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 542

CGGAARTAGA ACTGSGGACG GTAG

24

35

## 2) INFORMATION FOR SEQ ID NO: 543

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

45

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 543

50 ATCTTAGTAG TTTCTGCTGC TGA

23

55

## 2) INFORMATION FOR SEQ ID NO: 544

## (i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 544

5

AYGTTGTCGC CMGGCATTMC CAT

23

10 2) INFORMATION FOR SEQ ID NO: 545

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

(B) TYPE: Nucleic acid

15

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 545

TACATCCTBG TYGCICTIAA CAAGTG

26

25

2) INFORMATION FOR SEQ ID NO: 546

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

30

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 546

CCRCGICCGG TRATGGTGAA GAT

23

40

2) INFORMATION FOR SEQ ID NO: 547

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

45

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 547

GTACAGTTGC TTCAGGACGT ATC

23

55

2) INFORMATION FOR SEQ ID NO: 548

(i) SEQUENCE CHARACTERISTICS:

60



- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 548

10 ACGTTCGATT TCATCACGTT G

21

2) INFORMATION FOR SEQ ID NO: 549

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 549

25

GAACGTGATA CTGACAAACC TTTA

24

30 2) INFORMATION FOR SEQ ID NO: 550

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 550

GAAGAAGAAC ACCAACGTTG

20

45

2) INFORMATION FOR SEQ ID NO: 551

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 551

GAAGAAAAA TCTTCGAAC TGGCTA

25

60

## 2) INFORMATION FOR SEQ ID NO: 552

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 552

TACACGGCCG GTGACTACG

19

## 2) INFORMATION FOR SEQ ID NO: 553

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 553

GGCCGTGTTG AACGTGGTCA AATCA

25

## 2) INFORMATION FOR SEQ ID NO: 554

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 554

G TTCCTTACA TCGTTGTTTT TCTC

24

## 2) INFORMATION FOR SEQ ID NO: 555

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 555

TCTCGAACTT TCTCTATGTA TGCA

24

5

2) INFORMATION FOR SEQ ID NO: 556

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 556

CGGCGCNATC YTSGTTGTTG C

21

20

2) INFORMATION FOR SEQ ID NO: 557

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 557

35 CCMAGGCATR ACCATCTCGG TG

22

2) INFORMATION FOR SEQ ID NO: 558

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 558

50

TCITTYAART AYGCTGGGT

20

55 2) INFORMATION FOR SEQ ID NO: 559

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single

60

542

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 559

CCGACRGCRA YIGTYTGICK CAT

23

10

2) INFORMATION FOR SEQ ID NO: 560

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

15 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 560

GAYTTCATYA ARAAYATGAT YAC

23

25

2) INFORMATION FOR SEQ ID NO: 561

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

30 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 561

ACIGTICGGC CRCCCTCACG GAT

23

40

2) INFORMATION FOR SEQ ID NO: 562

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 562

55 CARATGRAYG ARCCICCIGG IGYIMGIATG

30

60

2) INFORMATION FOR SEQ ID NO: 563

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
5    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 563  
10 GGYTGRTAIC CIACIGCIGA IGGCAT 26

15 2) INFORMATION FOR SEQ ID NO: 564

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 29 bases  
    (B) TYPE: Nucleic acid  
20    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 564  
TAYGGICARA TGAAYGARCC ICCIGGIAA 29

30 2) INFORMATION FOR SEQ ID NO: 565

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 bases  
35    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 565  
GGYTGRTAIC CIACIGCIGA IGGDAT 26

45 2) INFORMATION FOR SEQ ID NO: 566

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 bases  
50    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

55 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 566  
TTYGGIGGIG CIGGIGTIGG IAARAC 26

60

## 2) INFORMATION FOR SEQ ID NO: 567

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 567  
15 TCRTCIGCIG GIACRTAIAY IGCYTG 26

## 2) INFORMATION FOR SEQ ID NO: 568

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 568  
30 RTIATIGGIG CIGTIRTIGA YGT 23

## 2) INFORMATION FOR SEQ ID NO: 569

35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 569  
RTIRTIGGIS CIGTIRTIGA TAT 23

## 2) INFORMATION FOR SEQ ID NO: 570

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
55 (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 570

RTIRYIGGIC CIGTIRTIGA YGT

23

5

2) INFORMATION FOR SEQ ID NO: 571

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

10

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 571

RTIRTIGGIC CIGTIRTIGA TGT

23

20

2) INFORMATION FOR SEQ ID NO: 572

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

25

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 572

35 RTIRTIGGIS CIGTIRTIGA

20

40

2) INFORMATION FOR SEQ ID NO: 573

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 573

50

CCICCIACCA TRTARAAIGC

20

55 2) INFORMATION FOR SEQ ID NO: 574

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

60

546

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 574

ATIGCIATGG AYGGIACIGA RGG 23

10

2) INFORMATION FOR SEQ ID NO: 575

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 575

TIACCATTTC AGTACCTTCT GGTA 25

25

2) INFORMATION FOR SEQ ID NO: 576

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 576

AACTTCRTCA AGAAGGTYGG TTACAA 26

40

2) INFORMATION FOR SEQ ID NO: 577

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 577

55 CATGATTGAA CCATCCACCA 20

2) INFORMATION FOR SEQ ID NO: 578

60



(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
5    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 578  
10 CATGATTGAA GCTTCCACCA 20

15 2) INFORMATION FOR SEQ ID NO: 579

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
20    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 579  
GAAGGCCGTG CTGGTGAGAA 20

30 2) INFORMATION FOR SEQ ID NO: 580

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
35    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 580  
GCTAAACCAG CTACAATCAC TCCAC 25

45 2) INFORMATION FOR SEQ ID NO: 581

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 bases  
50    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 581  
ACATCGGTGC ATTATTACGT GG 22

60

## 2) INFORMATION FOR SEQ ID NO: 582

5 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 28 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 582  
15 TTTCAACTTC GTCGTTGACA CGAACAGT 28

## 2) INFORMATION FOR SEQ ID NO: 583

20 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 35 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 583  
30 CAACTGCTTT TTGGATATCT TCTTTAATAC CAACG 35

## 2) INFORMATION FOR SEQ ID NO: 584

35 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 584  
ACATGACACA TCTAAAACAA 20

## 2) INFORMATION FOR SEQ ID NO: 585

50 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
55 (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
60 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 585

ACCACATACT GAATTCAAAG

20

5

2) INFORMATION FOR SEQ ID NO: 586

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 586

CAGAAGTATA CGTATTATCA

20

20

2) INFORMATION FOR SEQ ID NO: 587

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 587

35 CGTATTATCA AAAGACGAAG

20

2) INFORMATION FOR SEQ ID NO: 588

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 588

50

TCTTCTCAAA CTATCGTCCA

20

2) INFORMATION FOR SEQ ID NO: 589

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single

60

550

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 589

GCACGAAACT TCTAAAACAA

20

10

2) INFORMATION FOR SEQ ID NO: 590

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

15

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 590

TATACGTATT ATCTAAAGAT

20

25

2) INFORMATION FOR SEQ ID NO: 591

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 591

TCCTGGTTCT ATTACACCAC

20

40

2) INFORMATION FOR SEQ ID NO: 592

(i) SEQUENCE CHARACTERISTICS:

45

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 592

55

CAAAGCTGAA GTATACGTAT

20

60

2) INFORMATION FOR SEQ ID NO: 593

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 593

TTCACCTAACT ATCGCCCACA

20

2) INFORMATION FOR SEQ ID NO: 594

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 594

ATTGGTATCC ATGACACTTC

20

2) INFORMATION FOR SEQ ID NO: 595

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 595

TTAAAGCAGA CGTATACGTT

20

2) INFORMATION FOR SEQ ID NO: 596

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 596

GAAATTATTG GTATCAAAGA

20

## DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET  
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 2 DE 4

NOTE: Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

## JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE  
THAN ONE VOLUME

THIS IS VOLUME 2 OF 4

NOTE: For additional volumes please contact the Canadian Patent Office

## DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET  
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 3 DE 4

NOTE: Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

## JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE  
THAN ONE VOLUME

THIS IS VOLUME 3 OF 4

NOTE: For additional volumes please contact the Canadian Patent Office

## 2) INFORMATION FOR SEQ ID NO: 597

5 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 597  
15 ATTGGTATCA AAGAAACTTC 20

## 2) INFORMATION FOR SEQ ID NO: 598

20 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 598  
30 AATTACACCT CACACAAAAT 20

## 2) INFORMATION FOR SEQ ID NO: 599

35 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
    (B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 599  
CGGTGAAGAA ATCGAAATCA 20

## 2) INFORMATION FOR SEQ ID NO: 600

50 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
55 (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
60



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 600

ATGCAAGAAG AATCAAGCAA

20

5

2) INFORMATION FOR SEQ ID NO: 601

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

10

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 601

GTTTCACGTG ATGATGTACA

20

20

2) INFORMATION FOR SEQ ID NO: 602

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

25

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 602

35 AAGTTGAAGT TGTTGGTATT

20

2) INFORMATION FOR SEQ ID NO: 603

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

40

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 603

50

GGTATTAAAG ACGAAACATC

20

2) INFORMATION FOR SEQ ID NO: 604

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

60

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 604

GGTGATGAAG TAGAAATCGT

20

10

2) INFORMATION FOR SEQ ID NO: 605

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

15

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 605

GAAATGTTCC GTAAATTATT

20

25

2) INFORMATION FOR SEQ ID NO: 606

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

30

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 606

ATTAGACTAC GCTGAAGCTG

20

40

2) INFORMATION FOR SEQ ID NO: 607

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 821 bases

45

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecalis*

55

(B) STRAIN: ATCC 29212

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 607

CGGAGCTATC TTAGTAGTTT CTGCTGCTGA TGGTCCTATG CCTCAAACAC  
60 GTGAACATAT CTTATTATCA CGTAACGTTG GTGTACCATA CATCGTTGTA

50  
100

	TTCTTAAACA	AAATGGATAT	GGTTGATGAC	GAAGAATTAT	TAGAATTAGT	150
	AGAAATGGAA	GTTCGTGACT	TATTATCAGA	ATACGATTTC	CCAGGCGATG	200
	ATGTTCCAGT	TATCGCAGGT	TCTGCTTTGA	AAGCTTTAGA	AGGCGACGAG	250
	TCTTATGAAG	AAAAAATCTT	AGAATTAATG	GCTGCAGTTG	ACGAATATAT	300
5	CCCAACTCCA	GAACGTGATA	CTGACAAACC	ATTCATGATG	CCAGTCGAAG	350
	ACGTATTCTC	AATCACTGGA	CGTGGTACTG	TTGCTACAGG	ACGTGTTGAA	400
	CGTGGTGAAG	TTCGCGTTGG	TGACGAAGTT	GAAATCGTTG	GTATTAAAGA	450
	CGAAACATCT	AAAACAACCTG	TTACAGGTGT	TGAAATGTTT	CGTAAATTAT	500
	TAGACTACGC	TGAAGCAGGC	GACAACTTCG	GTGCTTTATT	ACGTGGTGTA	550
10	GCACGTGAAG	ATATCGAACG	TGGACAAGTA	TTAGCTAAAC	CAGCTACAAT	600
	CACCTCCACAC	ACAAAATTCA	AAGCTGAAGT	ATACGTATTA	TCAAAAGAAG	650
	AAGGCGGACG	TCACACTCCA	TTCTTCACTA	ACTACCGTCC	TCAATTCTAC	700
	TTCCGTACAA	CAGACGTTAC	TGGTGTTGTA	GAATTGCCAG	AAGGTACTGA	750
	AATGGTAATG	CCTGGTGATA	ACGTTGCTAT	GGACGTTGAA	TTAATTCACC	800
15	CAATCGCTAT	CGAAGACGGA	A			821

## 2) INFORMATION FOR SEQ ID NO: 608

20

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 751 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

25

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: ATCC 19434

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 608

35	CGGAGCTATC	TTGGTAGTTT	CTGCTGCTGA	CGGCCCAATG	CCTCAAACCTC	50
	GTGAACACAT	CCTATTGTCT	CGTCAAGTTG	GTGTTCCCTTA	CATCGTTGTA	100
	TTCTTGAACA	AAGTAGACAT	GGTTGATGAC	GAAGAATTAC	TAGAATTAGT	150
	TGAAATGGAA	GTTCGTGACC	TATTAACAGA	ATACRAATTC	CCTGGTGRCG	200
	ATGTTCTCTG	AGTTGCTGGA	TCAGCTTTGA	AAGCTCTAGA	AGGCGACGCT	250
40	TCATACGAAG	AAAAAATTCT	TGAATTAATG	GCTGCAGTTG	ACGAATACAT	300
	CCCAACTCCA	GAACGTGACA	ACGACAAACC	ATTCATGATG	CCAGTTGAAG	350
	ACGTGTTCTC	AATTACTGGA	CGTGGTACTG	TTGCTACAGG	TCGTGTTGAA	400
	CGTGGACAAG	TTCGCGTTGG	TGACGAAGTT	GAAGTTGTTG	GTATTGCTGA	450
	AGAAACTTCA	AAAACAACAG	TTACTGGTGT	TGAAATGTTT	CGTAAATTGT	500
45	TAGACYACGC	TGAAGCTGGA	GACRACATTG	GTGCTTTACT	ACGTGGTGTT	550
	GCACGTGAAG	ACATCCAACG	TGGACAAGTT	TTAGCTAAAC	CAGGTACAAT	600
	CACACCTCRT	ACAAAATTCT	CTGCAGAAAGT	ATACGTGTTG	ACAAAAGAAG	650
	AAGGTGGACG	TCATACTCCA	TTCTTCACTA	ACTACCGTCC	ACAATTCTAC	700
	TTCCGTACAA	CTGACGTAAC	AGGTGTTGTT	GAATTACCAG	AAGGAACTGA	750
50	A					751

## 2) INFORMATION FOR SEQ ID NO: 609

55

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 751 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Enterococcus gallinarum*  
(B) STRAIN: ATCC 49573

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 609

10	CGGTGCGATC	TTAGTAGTAT	CTGCTGCTGA	CGGTCCTATG	CCTCAAACCTC	50
	GTGAACACAT	CTTGTTATCA	CGTAACGTTG	GCGWACCATA	CATCGTTGTT	100
	TTCTTGAACA	AAATGGATAT	GGTTGAYGAC	GAAGAATTGC	TAGAATTAGT	150
	TGAAATGGAA	GTTTCGTGACC	TATTGTCTGA	ATATGACTTC	CCAGGCGACG	200
	ATGTTCCCTGT	AATCGCCGGT	TCTGCTTTGA	AAGCTCTTGA	AGGAGATCCT	250
15	TCATACGAAG	AAAAAATCAT	GGAATTGATG	GCTGCAGTTG	ACGAATACGT	300
	TCCAACCTCA	GAACGTGATA	CTGACAAACC	ATTCATGATG	CCAGTCGAAG	350
	ACGTATTCTC	AATCACTGGA	CGTGGTACTG	TTGCTACAGG	CCGTGTTGAA	400
	CGTGGACAAG	TTCGCGTTGG	TGATGAAGTA	GAAATCGTTG	GTATTGCTGA	450
	CGAAACTGCT	AAAACAACCTG	TAACAGGTGT	TGAAATGTTT	CGTAAATTGT	500
20	TAGACTATGC	TGAAGCAGGG	GATAACATTG	GTGCATTGCT	ACGTGGGGTT	550
	GCTCGTGAAG	ACATCCAACG	TGGACAAGTA	TTGGCTAAAG	CTGGTACRAT	600
	CACACCTCAT	ACAAAATTCA	AAGCTGAAGT	TTATGTTTTG	ACAAAAGAAG	650
	AAGGTGGRCG	TCACACTCCA	TTCTTCACTA	ACTACCGTCC	TCAGTTCCTAC	700
	TTCCGTACAA	CTGACGTAAC	TGGTGTGTGT	GAATTACCAG	AAGGAACTGA	750
25	A					751

2) INFORMATION FOR SEQ ID NO: 610

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 891 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
35 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 40 (A) ORGANISM: *Haemophilus influenzae* Rd  
(B) STRAIN: KW20  
(C) ACCESSION NUMBER: extracted from U32739

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 610

45	AATATGATTA	CTGGTGCGGC	ACAAATGGAT	GGTGCTATTT	TAGTAGTAGC	50
	AGCAACAGAT	GGTCCTATGC	CACAAACTCG	TGAACACATC	TTATTAGGTC	100
	GCCAAGTAGG	TGTTCCATAC	ATCATCGTAT	TCTTAAACAA	ATGCGACATG	150
	GTAGATGACG	AAGAGTTATT	AGAATTAGTC	GAAATGGAAG	TTCGTGAACT	200
50	TCTATCTCAA	TATGACTTCC	CAGGTGACGA	TACACCAATC	GTACGTGGTT	250
	CAGCATTACA	AGCGTTAAAC	GGCGTAGCAG	AATGGGAAGA	AAAAATCCTT	300
	GAGTTAGCAA	ACCACTTAGA	TACTTACATC	CCAGAACCAG	AACGTGCGAT	350
	TGACCAACCG	TTCTTCTTTC	CAATCGAAGA	TGTGTTCTCA	ATCTCAGGTC	400
	GTGGTACTGT	AGTAACAGGT	CGTGTAAGAC	GAGGTATTAT	CCGTACAGGT	450
55	GATGAAGTAG	AAATCGTCGG	TATCAAAGAT	ACAGCGAAAA	CTACTGTAAC	500
	GGGTGTTGAA	ATGTTCCGTA	AATTACTTGA	CGAAGGTCGT	GCAGGTGAAA	550
	ACATCGGTGC	ATTATTACGT	GGTACCAAAC	GTGAAGAAAT	CGAACGTGGT	600
	CAAGTATTAG	CGAAACCAGG	TTCAATCACA	CCACACACTG	ACTTCGAATC	650
	AGAAGTGTAC	GTATTATCAA	AAGATGAAGG	TGGTCGTCAT	ACTCCATTCT	700
60	TCAAAGGTTA	CCGTCCACAA	TTCTATTTCC	TACAACAGA	CGTGACTGGT	750

ACAATCGAAT	TACCAGAAGG	CGTGGAATG	GTAATGCCAG	GCGATAACAT	800
CAAGATGACA	GTAAGCTTAA	TCCACCCAAT	TGCGATGGAT	CAAGGTTTAC	850
GTTTCGCAAT	CCGTGAAGGT	GGCCGTACAG	TAGGTGCAGG	C	891

5

## 2) INFORMATION FOR SEQ ID NO: 611

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 818 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus epidermidis*  
 (B) STRAIN: ATCC 14990

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 611

CGGCGGTATC	TTAGTTGTAT	CTGCTGCTGA	CGGTCCAATG	CCACAAACTC	50
GTGAACACAT	CTTATTATCA	CGTAACGTTG	GTGTACCAGC	ATTAGTTGTA	100
TTCTTAAACA	AAGTTGACAT	GGTAGACGAC	GAAGAATTAT	TAGAATTAGT	150
TGAAATGGAA	GTTCGTGACT	TATTAAGCGA	ATATGACTTC	CCAGGTGACG	200
ATGTACCTGT	AATCGCTGGT	TCTGCATTAA	AAGCATTAGA	AGGCGATGCT	250
GAATACGAAC	AAAAAATCTT	AGACTTAATG	CAAGCAGTTG	ATGATTACAT	300
TCCAACCTCA	GAACGTGATT	CTGACAAACC	ATTCATGATG	CCAGTTGAGG	350
ACGTATTCTC	AATCACTGGT	CGTGGTACTG	TTGCTACAGG	CCGTGTTGAA	400
CGTGGTCAAA	TCAAAGTTGG	TGAAGAAGTT	GAAATCATCG	GTATGCACGA	450
AACTTCTAAA	ACAACGTGTA	CTGGTGTAGA	AATGTTCCGT	AAATTATTAG	500
ACTACGCTGA	AGCTGGTGAC	AACATCGGTG	CTTTATTACG	TGGTGTGCA	550
CGTGAAGACG	TACAACGTGG	TCAAGTATTA	GCTGCTCCTG	GTTCTATTAC	600
ACCACACACA	AAATTCAAAG	CTGAAGTATA	CGTATTATCT	AAAGATGAAG	650
GTGGACGTCA	CACTCCATTG	TTCACATACT	ATCGCCCAAC	ATTCTATTTC	700
CGTACTACTG	ACGTAACCTG	TGTTGTAAAC	TTACCAGAAG	GTACAGAAAT	750
GGTTATGCCT	GGCGACAACG	TTGAAATGAC	AGTTGAATTA	ATCGCTCCAA	800
TCGCTATCGA	AGACGGAA				818

40

## 2) INFORMATION FOR SEQ ID NO: 612

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 825 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*  
 serotype Paratyphi A  
 (B) STRAIN: ATCC 9150

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 612

GGCGCGATCC	TGGTTGTTGC	TGCGACTGAC	GGCCCGATGC	CGCAGACCCG	50
------------	------------	------------	------------	------------	----

558

	TGAGCACATC	CTGCTGGGTC	GTCAGGTAGG	CGTTCCGTAC	ATCATCGTGT	100
	TCCTGAACAA	ATGCGACATG	GTTGATGACG	AAGAGCTGCT	GGAAGTGGTT	150
	GAAATGGAAG	TTCGTGAACT	TCTGTCTCAG	TACGACTTCC	CGGGCGACGA	200
	CACGCCGATC	GTTCGTGGTT	CTGCTCTGAA	AGCGCTGGAA	GGCGACGCAG	250
5	AGTGGGAAGC	GAAAATCATC	GAAGTGGCTG	GCTTCCTGGA	TTCTTACATC	300
	CCGGAACCAG	AGCGTGCGAT	TGACAAGCCG	TTCTTGCTGC	CGATCGAAGA	350
	CGTATTCTCC	ATCTCCGGTC	GTGGTACCGT	TGTTACCGGT	CGTGAGAAC	400
	GCGGTATCAT	CAAAGTGGGC	GAAGAAGTTG	AAATCGTTGG	TATCAAAGAG	450
	ACTCAGAAGT	CTACCTGTAC	TGGCGTTGAA	ATGTTCCGCA	AACTGCTGGA	500
10	CGAAGGCCGT	GCCGGTGAGA	ACGTAGGTGT	TCTGCTGCGT	GGTATCAAAC	550
	GTGAAGAAAT	CGAACGTGGT	CAGGTACTGG	CTAAGCCGGG	CACCATCAAG	600
	CCGCACACCA	AGTTCGAATC	TGAAGTGATC	ATTCTGTCCA	AAGATGAAGG	650
	CGGCCGTCAT	ACTCCGTTCT	TCAAAGGCTA	CCGTCCGCAG	TTCTACTTCC	700
	GTACTACTGA	CGTGACTGGC	ACCATCGAAC	TGCCGGAAGG	CGTAGAGATG	750
15	GTAATGCCGG	GCGACAACAT	CAAAATGGTT	GTTACCCTGA	TCCACCCGAT	800
	CGCAATGGAC	GACGGTCTGC	GTTTC			825

## 20 2) INFORMATION FOR SEQ ID NO: 613

## (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 778 bases
	(B) TYPE: Nucleic acid
25	(C) STRANDEDNESS: Double
	(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

	(A) ORGANISM: <i>Serratia ficaria</i>
	(B) STRAIN: ATCC 33105

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 613

35	GGCGCTATCC	TGGTTGTTGC	TGCGACTGAC	GGCCCAATGC	CTCAGACCCG	50
	TGAGCACATC	CTGCTGGGTC	GYCAGGTTGG	CGTTCCTTTC	ATCATCGTRT	100
	TCATGAACAA	ATGCGACATG	GTTGATGATG	AAGAGCTGCT	GGAAGTGGTA	150
	GAAATGGAAG	TTCGCGAACT	GCTGTCCGCT	TACGACTTCC	CTGGCGATGA	200
40	CCTGCCGGTG	ATTGCGGGTT	CCGCGCTGAA	AGCGCTGGAA	GGCGAAGCCG	250
	AGTGGGAAGC	TAAAATCATC	GAGCTGGCTG	AMCMSCTGGA	TACTTACATC	300
	CCAGAACCAG	AGCGCGCTAT	CGACAAGCCG	TTCTTGCTGC	CAATCGAAGA	350
	CGTATTCTCC	ATCTCCGGTC	GTGGTACCGT	GTTTACCGGT	CGTGTTGAGC	400
	GCGGTATCAT	CAAAGTTGGC	GAAGAAGTTG	AAATCGTTGG	TATCAAAGAC	450
45	ACCGTCAAGT	CTACCTGTAC	CGGCGTTGAA	ATGTTCCGCA	AACTGCTGGA	500
	CGAAGGCCGT	GCCGGTGAGA	ACGTAGGTGT	TCTGCTGCGT	GGTATCAAGC	550
	GTGAAGACAT	CGAACGTGGT	CAGGTTCTGG	CTAAACCAGG	TTCCATCAAG	600
	CCGCACACCC	AGTTCGATTC	AGAAGTGATC	ATCCTGAGCA	AAGAAGAAGG	650
	TGGTCGTCAC	ACKCCATTCT	TCAAAGGCTA	CCGTCCACAG	TTCTACTTCC	700
50	GTACTACTGA	CGTGACCGGT	ACCATCGAAC	TGCCAGAAGG	CGTTGAGATG	750
	GTAATGCCTG	GCGACAACGT	GAACATGA			778

## 55 2) INFORMATION FOR SEQ ID NO: 614

## (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 653 bases
	(B) TYPE: Nucleic acid
60	(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus malodoratus*

(B) STRAIN: ATCC 43197

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 614

```

10 GTGCGATCTT AGTAGTATCA GCTACTGATG GTCCAATGCC TCAAACCTCGT      50
   GAACACATTT TGTTATCACG TCAAGTTGGT GTTAAGCACT TGATCGTTTT      100
   CTTGAACAAA GTAGATTTAG TTGATGACGA AGAATTGATC GACTTAGTTG      150
   AAATGGAAGT ACGTGAATTA CTTTCTGAAT ATGGTTTCCC AGGTGATGAT      200
15 ATTCCAGTGC TTAAAGGTTT TGCTTTGAAA GCATTAGAAG GCGATCCAGA      250
   ACAAGAACAA GTTATTCTTG ATTTGATGGA TACCGTTGAT GAATATATCC      300
   CAACACCTGA ACGTGACAAT GACAAACCGT TCTTGTTACC AGTTGAGGAT      350
   GTTTTCTCGA TCACAGGACG TGGTACTGTA GCTTCTGGTC GTATCGACCG      400
   TGGCGAAGTT AAAGTCGGCG ATGAAATTGA AATCATCGGG ATCAAACCTG      450
20 AAGTTCAAAA AGCAATCGTT ACTGGACTTG AAATGTTCCG TAAACATTG      500
   GATTATGGTG AAGCTGGCGA TAACGTTGGG GTTCTATTAC GTGGGATTAC      550
   ACGTGATGAA ATCGAACGTG GCCAAGTATT AGCTAAACCA GGTTCAATCA      600
   CACCACATAC TAAGTTCAAA GCCGAAGTAT ATGTGTTGAC GAAAGAAGAA      650
   GGT
25

```

2) INFORMATION FOR SEQ ID NO: 615

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 829 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus durans*

40 (B) STRAIN: ATCC 19432

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 615

```

45 CCATTCTAGT TGTATCTGCA ACAGATGGAC CAATGCCACA AACACGTGAA      50
   CATATTTTAT TGTCACGTCA AGTAGGTGTT AAATATTTGA TCGTCTTCTT      100
   GAACAAAATC GACTTAGTAG ATGATGAAGA ATTGATTGAT CTTGTGCGAA      150
   TGGGAAGTTCG TGAATTATTA AGCGAATATG GTTCCCAGG TGACGATACA      200
   CCAGTCATCA AAGGTTTCAGC ATTAAGAGCT TTACAAGGAG ATCCTGATGC      250
   AGAAGCAGCT ATCATGGAAT TGATGGATAC TGTTGATGAA TATATCCCAA      300
50 CACCAGAACG TGATACAGAC AAACCATTAT TGTTACCAGT GGAAGATGTC      350
   TTCTCAATCA CAGGTCGTGG GACTGTTGCT TCAGGTCGTA TCGATCGTGG      400
   TGCAGTTCGT GTAGGTGATG AAATCGAAAT CGTCGGTATC AAACCTGAAA      450
   CACAAAAGC TGTGTAACT GGGGTCGAAA TGTTCCGCAA GACATTAGAC      500
   TATGGTGAAG CAGGAGATAA CGTTGGGGTA TTGTTACGTG GTATCCAACG      550
55 TGAAGATATC GAACGTGGAC AAGTAATCGC AAAACCAGGT TCAATCACAC      600
   CACATACAAA ATTCAAAGCA GAAGTGACG TATTGACAAA AGAAGAAGGT      650
   GGACGTCATA CACCATTCTT CAATAACTAT CGTCCACAAT TCTACTCCG      700
   TACAACCTGAC GTAACCTGAA CAATCGTTTT ACCTGGAGGC ACTGAAATGG      750
   TTATGCCTGG AGATAACGTA ACGATCGACG TTGAATTGAT CCATCCAGTT      800
60 GCCATCGAAA ACGGAACAAC TTTCTCTAT      829

```

## 2) INFORMATION FOR SEQ ID NO: 616

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 669 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus pseudoavium*  
 (B) STRAIN: ATCC 49372

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 616

20	GGTGCAATTT TAGTAGTATC TGCTACTGAT GGCCCAATGC CACAAACACG	50
	TGAACATATC TTGTTATCAC GTCAAGTAGG GGTTAAACAC TTAATCGTCT	100
	TCTTGAACAA AGTTGATTTA GTTGATGATG AAGAATTGAT CGATTTAGTT	150
	GAAATGGAAG TTCGGAATT GCTTCTGAA TATGGTTTCC CAGGCGATGA	200
	TATTCCAGTA CTTAAAGGTT CTGCTTTGAA AGCTTTAGAA GGCGATCCTG	250
25	AACAAGAACA AGTAATCCTT GACTTGATGG ATACGGTTGA TGAATACATC	300
	CCAACGCCTG AACGTGATAC TGACAAACCA TTCTTGTTAC CAGTCGAAGA	350
	TGTCTTCTCA ATCACAGGAC GTGGTACGGT TGCATCTGGT CGTATCGATC	400
	GTGGGGAAGT TAAAGTCGGT GATGAAGTTG AAATCATCGG GATCAAACCT	450
	GAAGTGCAAA AAGCTGTCGT AACTGGACTA GAAATGTTCC GTAAGACATT	500
30	GGATTACGGT GAAGCTGGCG ATAACGTTGG GGTTCATTA CGTGGGATTA	550
	CTCGTGATGA AATCGAACGT GGACAAGTAT TAGCTAAACC AGGTTCAATC	600
	ACTCCACATA CGAAATTCAG TGCAGAAGTT TATGTATTGA CGAAAGAAGA	650
	AGGTGGCCGT CATACGCCA	669

## 2) INFORMATION FOR SEQ ID NO: 617

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 835 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus dispar*  
 (B) STRAIN: ATCC 51266

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 617

	CGGGGCAATT TTAGTTGTAT CTGCAACTGA TGGCCCAATG CCACAAACAC	50
	GTGAACACAT TTTGTTAGCT CGTCAAGTAG GGGTTAAATA TTAAATCGTC	100
55	TTCTTGAACA AAACAGATTT AGTTGATGAT GAAGAATTAT TGGAAGTAGT	150
	TGAAATGGAA GTTCGTGAAT TATTAAATGA ATACAATTTC CCTGGCGATG	200
	ATATTCCTGT TATTCGCGGA TCTGCTTTAA AAGCATTAGA AGGCGATCCA	250
	GAACAAGAAG AAGTAATTAT GAACTTGATG GATACTGTGG ATGAATATAT	300
	CCCAACTCCA GAACGTGACA ATGATAAACC ATTCTTGTTA CCAGTGGAAG	350
60	ATGTCTTCAC AATTACTGGT CGTGGTACTG TTGCTTCAGG TCGTATCGAC	400



	CGTGGTAAAG	TCAACGTTGG	TGATGAAATT	GAAATTATCG	GAATTAAACC	450
	AGAAACACAA	AAAGCTGTTG	TAACCGGTTT	GGAAATGTTT	CGTAAAACTT	500
	TGGATTATGG	TGAAGCTGGT	GATAACGTTG	GGGTCTTATT	ACGTGGGATT	550
	ACTCGTGATG	AAGTAGAACG	TGGTCAAGTA	TTAGCAAAAC	CAGGTTCAT	600
5	TACACCGCAT	ACCAAATTTA	AAGGTGAAGT	TTATATCTTA	ACAAAAGAAG	650
	AAGGTGGACG	TCATACTCCT	TTCTTTAATA	ACTATCGTCC	TCAATTTTAT	700
	TTCCGTACAA	CTGATGTGAC	TGGTAACATC	GCATTACCTG	AAGGAACTGA	750
	AATGGTAATG	CCTGGTGATA	ATGTAACAAT	TGAAGTTGAA	TTGATTCATC	800
	CAATCGCCGT	TGAAAAAGGG	ACTACTTTCT	CAATT		835

10

## 2) INFORMATION FOR SEQ ID NO: 618

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 673 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 20 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Enterococcus avium*  
 (B) STRAIN: ATCC 14025
- 25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 618

	GGTGCAATCC	TAGTAGTATC	AGCTACTGAT	GGTCCAATGC	CGCAAACACG	50
30	TGAACATATT	TTGCTATCAC	GGCAAGTGGG	TGTTAAACAC	TTAATCGTAT	100
	TTTTAAACAA	AGTTGATTTA	GTCGATGATG	AAGAAATTGAT	CGATCTAGTT	150
	GAAATGGAAG	TCCGTGAATT	ACTTTCTGAA	TATGGTTTCC	CAGGTGACGA	200
	TATTCCAGTT	CTCAAAGGTT	CAGCTTTGAA	AGCATTAGAA	GGCGATCCTG	250
	AACAAGAACA	AGTAATCCTT	GATTTAATGG	ATACAGTTGA	CGAATATATC	300
35	CCAACTCCAG	AACGTGACAC	TGACAAGCCA	TTCTTGTTAC	CAGTCGAAGA	350
	TGTATTTTCT	ATCACTGGTC	GTGGGACTGT	AGCGTCTGGA	CGGATTGATC	400
	GTGGTGAAGT	TAAAGTCGGC	GATGAAGTTG	AAATCATCGG	GATCAAACCT	450
	GAAATTCAAA	AAGCAGTCGT	AACTGGACTT	GAAATGTTCC	GTAAAACCTT	500
	AGATTATGGT	GAAGCTGGCG	ATAACGTTGG	GGTCTATTA	CGTGGGATTA	550
40	CACGTGATGA	AATCGAACGT	GGTCAAGTCT	TAGCTAAACC	AGGTTCAATC	600
	ACACCACATA	CAAAATTCAG	TGCAGAAGTT	TACGTATTGA	CGAAAGAAGA	650
	AGGTGGACGT	CATACACCAT	CTT			673

45

## 2) INFORMATION FOR SEQ ID NO: 619

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1713 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- 55 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Saccharomyces cerevisiae*  
 (C) ACCESSION NUMBER: K00428
- 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 619

```

TTTCATATTT TTAAGGATTT TGTTTTAGCA CCCATCCGAC CTCAGTCAAT      50
ATATCCTTTT GCGACCAGGC TTTCCCTCCCT TTTGCTGCTA ACTGGTTACA      100
GATTTTCCCTA TTTTGGGTCA TTTTATCTT TGAAACTGAT TAAGCTGAAA      150
5  AAATTTGAGC TTCTTTGTTG TAAACTATTT TGTGCTTTCA GTTTTATTCT      200
AGCTCGACAA AGGTAACAGA CAAAATGTC AGCTTTATTA CCAAGATTAC      250
TCACAAGAAC AGCTTTTAAA GCTTCTGGGA AACTTCTGAG GCTCTCTTCA      300
GTAATTTCTA GGACCTTTTC TCAAACACT ACTTCCTATG CAGCTGCTTT      350
10 TGATCGTTCC AAACCGCATG TAAATATAGG TACGATCGGC CATGTTGATC      400
ATGGGAAGAC AACTTTAACC GCAGCCATTA CGAAAACGTT AGCCGCAAAA      450
GGTGGTGCCA ACTTCTTGGA CTATGCTGCC ATCGATAAGG CTCCGGAAGA      500
AAGAGCTCGT GGTATTACAA TTTCTACTGC ACACGTGGAA TACGAAACGG      550
CCAAGAGACA TTATTCTCAC GTCGACTGTC CAGGCCACGC TGATTACATC      600
AAGAATATGA TTACCGGTGC TGCTCAAATG GATGGTGCTA TCATTGTTGT      650
15 AGCTGCTACC GATGGACAAA TGCCCCAAAC TAGAGAACAT TTTACTTTGG      700
CCAGACAAGT TGGTGTCCAA CATATTGTCG TTTTGTGTTA CAAGGTTGAT      750
ACCATTGATG ATCCAGAAAT GTTAGAGTTA GTCGAAATGG AAATGAGAGA      800
ACTTTTAAAC GAATATGGGT TTGACGGTGA TAATGCTCCA ATTATCATGG      850
GTTCTGCCCT TTGCGCTTTG GAAGGTCGCC AACCTGAAAT TGGGGAGCAG      900
20 GCCATCATGA AACTTTTGGG TGCACTGGAT GAGTATATTC CTACACCTGA      950
AAGAGATTTG AACAAGCCTT TCTTGATGCC CGTTGAAGAT ATCTTCTCTA     1000
TCTCCGGTAG AGGTACTGTG GTCACTGGTC GTGTGGAAAG GGGTAATTTA     1050
AAGAAAGGTG AGGAATTGGA AATTGTTGGT CACAACCTCA CCCCATTGAA     1100
AACACAGTT ACTGGTATTG AAATGTTTAG AAAGGAATTG GACTCTGCTA     1150
25 TGGCAGGTGA CAATGCCGGT GTTTTACTTA GAGGTATCAG GAGAGATCAA     1200
TTGAAGAGAG GTATGGTCTT AGCTAAGCCA GGTACCGTTA AAGCCCATAC     1250
AAAGATTCTA GCCTCTTTGT ACATTTTATC CAAAGAGGAA GGTGGTAGAC     1300
ATTCTGGGTT TGGTGAAAAC TACAGACCAC AAATGTTTAT AAGAACAGCT     1350
GATGTTACAG TTGTGATGAG ATTTCCCTAAG GAGGTTGAAG ATCATTCTAT     1400
30 GCAAGTTATG CCAGGTGACA ATGTTGAAAT GGAATGTGAT TTGATCCATC     1450
CTACCCCATT AGAAGTTGGT CAACGTTTCA ATATCAGAGA GGGTGGAAGA     1500
ACTGTTGGTA CCGGTCTAAT CACACGTATT ATTGAATAGA CTTATTGATG     1550
CAACTGGAGT ATATTTCTAT ATATTCTGTT CATTTCCCCT CTCATAATAT     1600
ATACTTGTTT CGTTAAAATT TTATACGTGT AAATAAAGTG CCATAAATTT     1650
35 TTCAGCTTTA CTTTTGGTAG AGTCCTGCTA GCACTAGATT TTACAATTTT     1700
ATGTGCACAC ACC                                         1713

```

40 2) INFORMATION FOR SEQ ID NO: 620

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 18 bases
(B) TYPE: Nucleic acid
45 (C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: DNA

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 620

ATTGGTGCAT TGCTACGT

18

55

2) INFORMATION FOR SEQ ID NO: 621

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 751 bases
60 (B) TYPE: Nucleic acid

```

563

(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*  
(B) STRAIN: ATCC 19434

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 621

	TGGTGCAATC	TTAGTTGTTT	CTGCAACTGA	CGGTCCGATG	CCTCAAACAC	50
	GTGAACACAT	TTTATTGTCA	CGCCAAGTTG	GTGTAAAATA	CCTGATTGTT	100
	TTCTTGAACA	AAGTTGATTT	AGTCGATGAT	GAAGAATTGA	TCGATTGTTG	150
15	AGAAATGGAA	GTTCGCGAGT	TATTGAGCGA	ATATGGTTTC	CCAGGCGATG	200
	ACACTCCTGT	GATCAAAGGT	TCCGCATTAA	AAGCATTGCA	AGGCGATCCA	250
	GATGCTGAAG	CTGCTATTAT	GGAATTGATG	GATACAGTAG	ATGAATATAT	300
	CCCAACACCA	GAACGTGATA	CAGATAAACC	ATTACTATTG	CCAGTGGAAG	350
	ACGTCTTCTC	AATTACAGGT	CGAGGAACTG	TTGCCTCAGG	TCGTATTGAT	400
20	CGTGGTGCTG	TTCGTGTCGG	TGATGAGGTA	GAGATCGTAG	GGATCAAACC	450
	TGAAACACAA	AAAGCAGTTG	TAACAGGTGT	AGAAATGTTT	CGTAAAACGT	500
	TAGATTACGG	GGAAGCTGGG	GATAACGTAG	GCGTGTTGTT	ACGGGGGATC	550
	CAACGTGACG	ATATCGAACG	TGGACAAGTA	CTTGCTAAAC	CAGGTTCCAT	600
	TACTCCACAT	ACAAAATTCA	AAGCAGAAGT	GTACGTGTTG	ACAAAAGAAG	650
25	AAGGTGGACG	TCATACTCCA	TTCTTCAACA	ACTATCGTCC	ACAGTTCTAC	700
	TTCCGCACAA	CTGATGTTAC	AGGAACAATC	ACATTGCCAG	AAGATACAGA	750
	A					751

30

-2) INFORMATION FOR SEQ ID NO:- 622 -

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 750 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*  
(B) STRAIN: ATCC 13264

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 622

	GTCAAATGGG	ACGAATCCAG	ATTCCAAGAA	ATTGTCAAGG	AAACCTCCAA	50
	CTTTATCAAG	AAGGTTGGTT	ACAACCCAAA	GACTGTTCCA	TTCGTCCCAA	100
	TCTCTGGTTG	GAACGGTGAC	AACATGATTG	AAGCTACCAC	CAACGCTCCA	150
50	TGGTACAAGG	GTTGGGAAAA	GGAAACCAAG	GCCGGTGTCT	TCAAGGGTAA	200
	GACTTTGTTG	GAAGCCATTG	ACGCCATTGA	ACAACCATCT	AGACCAACTG	250
	ACAAGCCATT	GAGATTGCCA	TTGCAAGATG	TTTACAAGAT	TGGTGGTATT	300
	GGTACTGTGC	CAGTCGGTAG	AGTTGAAACC	GGTGTCTATC	AGCCAGGTAT	350
	GGTTGTTACT	TTCGCCCCAG	CTGGTGTTAC	CACTGAAGTC	AAGTCCGTTG	400
55	AAATGCATCA	CGAACAATTG	GAACAAGGTG	TTCCAGGTGA	CAACGTTGGT	450
	TTCAACGTCA	AGAACGTTTC	CGTTAAGGAA	ATCAGAAGAG	GTAACGTCTG	500
	TGGTGACGCT	AAGAACGATC	CACCAAAGGG	TTGCGCTTCT	TTCAACGCTA	550
	CCGTCATTGT	TTTGAACCAT	CCAGGTCAAA	TCTCTGCTGG	TTACTCTCCA	600
	GTTTTGGATT	GTCACACTGC	TCACATTGCT	TGTAGATTCT	ACGAATTGTT	650
60	GGAAAAGAAC	GACAGAAGAT	CTGGTAAGAA	GTTGGAAGAC	CATCCAAAGT	700

TCTTGAAGTC CGGTGACGCT GCTTTGGTCA AGTTCGTTCC ATCTAAGCCA

750

## 5 2) INFORMATION FOR SEQ ID NO:623

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1269 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptococcus neoformans*  
 (B) STRAIN: ATCC 44104

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 623

20 TCTTGAAAGC TTAAGGCCGA GCGAGAGCGA GGTATCACCA TCGACATTGC 50  
 TCTTTGGAAG TTCGAGACCC CCAGGTACCA GGTACCGTC ATTGACGCCC 100  
 CCGGTCACCG AGACTTCATC AAGAACATGA TCACCGGTAC CTCCCAGGCT 150  
 GACTGTGCCA TCCTCATCAT TGCCACCGGT ATCGGTGAGT TCGAGGCCGG 200  
 25 TATCTCCAAG GACGGTCAGA CCCGAGAGCA CGCCCTCCTC GCCTTCACCC 250  
 TCGGTGTCAG GCAGCTCATT GTTGCTTGCA ACAAGATGGA CACCTGCAAG 300  
 TGGTCCGAGG ACCGATTCAA CGAAATCGTC AAGGAGACCA ACGGTTTCAT 350  
 CAAGAAGGTT GGTTACAACC CCAAGGCTGT CCCCTTCGTC CCCATCTCTG 400  
 GTTGGCACGG TGACAACATG TTGGAGGAGA CCACCAAGTC AGTGAATCCG 450  
 30 CTTCTACGTG ATGAGATGTT TTTCTGACTT TCCGTGCAGC ATGCCCTGGT 500  
 ACAAGGGATG GACCAAGGAG ACCAAGTCTG GTGTTTCCAA GGGTAAGACC 550  
 CTCCTCGAGG CCATCGACGC CATCGAGCCC CCTACCCGAC CCACCGACAA 600  
 GCCCCTCCGT CTCCCTCTCC AGGACGTCTA CAAGATCGGT GGTATCGGCA 650  
 CAGTCCCTGT CGGCCGAGTC GAGACCGGTG TCATCAAGGC CGGTATGTTG 700  
 35 TCTCATCTCT CTGTCTCTGT AACATGCGTC TCGTAACATG CGCTTACTTC 750  
 ATTTTCAGGT ATGGTCGTCA AGTTCGCCCC CACCAACGTC ACCACTGAAG 800  
 TCAAGTCCGT TGAGATGCAC CACGAGCAGA TCCCGAGGG TCTTCCCGGA 850  
 GACAACGTTG GTTTCAACGT CAAGAACGTT TCCATCAAGG ACATCCGACG 900  
 AGGTAACGTC TGTGGTGACT CCAAGAACGA CCCCCTATG GAGGCTGCTT 950  
 40 CTTTCAACGC CCAGGTTATC GTCCTTAACC ACCCTGGTCA GATCGGTGCC 1000  
 GGTTACACCC CCGTTCTCGA CTGTCACT GCGCGTAAGC CTGACCCAAT 1050  
 ACCTCCAACA TACCTTTGAA GCTGACCCTT TCTAGACATT GCCTGCAAGT 1100  
 TTGCTGAGTT GATCGAGAAG ATTGACCGAC GAACCGGTAA GGTCATGGAG 1150  
 GCCGCCCCCA AGTTCGTCAA GTCTGGTGAC GCCGCCATTG TCAAGCTTGT 1200  
 45 TGCCCAAGG CCCCTCTGTG TTGAGACCTA CGCCGACTAC CCCCCTCTTG 1250  
 GTCGATTCGC CGTCCGAGA 1269

## 50 2) INFORMATION FOR SEQ ID NO: 624

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 753  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 60 (vi) ORIGINAL SOURCE:

565

(A) ORGANISM: *Candida albicans*  
 (B) STRAIN: ATCC 36801

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 624

```

5  TCTGTCAAAT GGGACAAAAA CAGATTTGAA GAAATCATCA AGGAAACCTC      50
   CAACTTCGTC AAGAAGGTTG GTTACAACCC AAAGACTGTT CCATTCGTTC      100
   CAATCTCTGG TTGGAATGGT GACAACATGA TTGAACCATC CACCAACTGT      150
   CCATGGTACA AGGGTTGGGA AAAGGAAACC AAATCCGGTA AAGTTACTGG      200
10  TAAGACCTTG TTAGAAGCTA TTGACGCTAT TGAACCACCA ACCAGACCAA      250
   ATTGACAAACC ATTGAGATTG CCATTGCAAG ATGTTTACAA GATCGGTGGT      300
   ATTGGTACTG TGCCAGTCGG TAGAGTTGAA ACTGGTATCA TCAAAGCCGG      350
   TATGGTTGTT ACTTTCGCCC CAGCTGGTGT TACCACTGAA GTCAAATCCG      400
   TTGAAATGCA TCACGAACAA TTGGCTGAAG GTGTTCCAGG TGACAATGTT      450
15  GGTTCACACG TTAAGAACGT TTCCGTTAAA GAAATTAGAA GAGGTAACGT      500
   TTGTGGTGAC TCCAAGAACG ATCCACCAAA GGGTTGTGAC TCTTTCAATG      550
   CCCAAGTCAT TGTTTTGAAC CATCCAGGTC AAATCTCTGC TGGTTACTCT      600
   CCAGTCTTGG ATTGTCACCC TGCCCACATT GCTTGTAAT TCGACACTTT      650
   GGTGAAAAG ATTGACAGAA GAACTGGTAA GAAATTGGAA GAAAATCCAA      700
20  AATTCGTCAA ATCCGGTGAT GCTGCTATCG TCAAGATGGT CCCAACCAAA      750
   CCA
  
```

25 2) INFORMATION FOR SEQ ID NO: 625

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus agalactiae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 625

```

40  CGTTGAAGAC ACGACCCAAA GTATCC      26
  
```

2) INFORMATION FOR SEQ ID NO: 626

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus agalactiae*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 626

```

60  TACCACCTTT TAAGTAAGGT GCTAAT      26
  
```

## 2) INFORMATION FOR SEQ ID NO: 627

- 5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear
- 10 (ii) MOLECULE TYPE: DNA
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Streptococcus agalactiae*
- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 627

ATTGTCTATA AAAATGGCGA TAAGTC

26

## 2) INFORMATION FOR SEQ ID NO: 628

- 25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear
- 30 (ii) MOLECULE TYPE: DNA
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Streptococcus agalactiae*
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 628

AAAATGGCGA TAAGTCACAA AAAGTA

26

## 2) INFORMATION FOR SEQ ID NO: 629

- 45 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear
- 50 (ii) MOLECULE TYPE: DNA
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Chlamydia pneumoniae*
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 629

AAGTTCCATC TCAACAAGGT CAATA

25

## 2) INFORMATION FOR SEQ ID NO: 630

567

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 bases  
    (B) TYPE: Nucleic acid  
5      (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear  
  
    (ii) MOLECULE TYPE: DNA  
  
10     (vi) ORIGINAL SOURCE:  
        (A) ORGANISM: *Chlamydia pneumoniae*  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 630  
15     CGGAGCTATC CTAGTCGTTT CA 22

20     2) INFORMATION FOR SEQ ID NO: 631  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 26 bases  
        (B) TYPE: Nucleic acid  
        (C) STRANDEDNESS: Single  
25      (D) TOPOLOGY: Linear  
  
    (ii) MOLECULE TYPE: DNA  
  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 631  
30     CAGACCAACY--GAIAARCCAT TRAGAT 26

35     2) INFORMATION FOR SEQ ID NO: 632  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 23 bases  
        (B) TYPE: Nucleic acid  
40      (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear  
  
    (ii) MOLECULE TYPE: DNA  
  
45     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 632  
    CCCTTTGGTG GRTCSTKCTT GGA 23

50     2) INFORMATION FOR SEQ ID NO: 633  
    (i) SEQUENCE CHARACTERISTICS:  
        (A) LENGTH: 26 bases  
55      (B) TYPE: Nucleic acid  
        (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear  
  
    (ii) MOLECULE TYPE: DNA  
60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 633

CAGACCAACY GAIAARCCIT TRAGAT

26

5

2) INFORMATION FOR SEQ ID NO: 634

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

10

(ii) MOLECULE TYPE: DNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 634

AACACYGTCA GRRCIATTGC YATGGA

26

20

2) INFORMATION FOR SEQ ID NO: 635

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

25

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 635

AAACCRGTIA RRGCRCTCT IGCTCT

26

35

2) INFORMATION FOR SEQ ID NO: 636

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

40

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 636

ACTGGYGTTG AIATGTTCCG YAA

23

50

2) INFORMATION FOR SEQ ID NO: 637

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

55

60



(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 637

ACGTCAGTIG TACGGAARTA GAA 23

10

2) INFORMATION FOR SEQ ID NO: 638

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 638

CCAATGCCAC AAACICGTGA RCACAT 26

25

2) INFORMATION FOR SEQ ID NO: 639

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 28 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 639

TTTACGGAAC ATTCWACAC CWGTIACA 28

40

2) INFORMATION FOR SEQ ID NO: 640

45 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 640

55 TCCATGGTIT WYGGICARAT GAA 23

60 2) INFORMATION FOR SEQ ID NO: 641

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 641

TGATAACCWA CIGCIGAIGG CATACG

26

## 2) INFORMATION FOR SEQ ID NO: 642

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 642

GGCGTIGGIG ARCGIACICG TGA

23

## 2) INFORMATION FOR SEQ ID NO: 643

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 643

ACTGGIGTIG ARATGTTCCG YAA

23

## 2) INFORMATION FOR SEQ ID NO: 644

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 644

ACGTCIGTIG TICKGAARTA GAA

23

## 2) INFORMATION FOR SEQ ID NO: 645

5 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 23 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 645  
15 ACGTCIGTIG TICKGAARTA RAA 23

## 2) INFORMATION FOR SEQ ID NO: 646

20 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 23 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 646  
30 ATCGACAAGC CITTCYTIAT GSC 23

## 2) INFORMATION FOR SEQ ID NO: 647

35 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 bases  
    (B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 647  
ACGTCCGTSG TRCGGAAGTA GAACTG 26

## 2) INFORMATION FOR SEQ ID NO: 648

50 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 26 bases  
55 (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 648

ACGTCSGTSG TRCGGAAGTA GAACTG

26

5

2) INFORMATION FOR SEQ ID NO: 649

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 649

20 GTCCTATGCC TCARACWCGI GAGCAC

26

2) INFORMATION FOR SEQ ID NO: 650

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
30 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 650

35 TTACGGAACA TYTCAACACC IGT

23

2) INFORMATION FOR SEQ ID NO: 651

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
45 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 651

50

TGACGACCAC CITCYTCYTT YTTCA

25

55 2) INFORMATION FOR SEQ ID NO: 652

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 bases  
(B) TYPE: Nucleic acid  
60 (C) STRANDEDNESS: Single

573

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 652

CCWAYAGTIY KICCCICYTC YCTIATA

27

10

2) INFORMATION FOR SEQ ID NO: 653

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

15

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 653

GAYTTCATIA ARAAYATGAT

20

25

2) INFORMATION FOR SEQ ID NO: 654

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

30

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 654

TACAARATYK GIGGTATYGG

20

40

2) INFORMATION FOR SEQ ID NO: 655

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

45

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 655

CCRATACCIC MRATYTTGTA

20

55

2) INFORMATION FOR SEQ ID NO: 656

60

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 656

AATTAATGGC TGCAGTTGAY GA

22

2) INFORMATION FOR SEQ ID NO: 657

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 657

TTGTCCACGT TCGATRTCTT CA

22

2) INFORMATION FOR SEQ ID NO: 658

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 658

GATYTAGTCG ATGATGAAGA ATT

23

2) INFORMATION FOR SEQ ID NO: 659

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 659

GCTTTTTGIG TTTCWGGTTT RAT

23

## 2) INFORMATION FOR SEQ ID NO: 660

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 660

15 GTAGAATTGA GGACGGTAGT TAG 23

## 2) INFORMATION FOR SEQ ID NO: 661

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 25 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 661

30 GTAGAAATGT GGWCGATART TRT 23

## 35 2) INFORMATION FOR SEQ ID NO: 662

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 832 bases  
 (B) TYPE: Nucleic acid  
 40 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

45 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Corynebacterium diphtheriae*  
 (B) STRAIN: ATCC 27010

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 662

50 CGGCGCAATC CTCGTTGTTG CTGCCACCGA CGGCCCAATG CCTCAGACCC 50  
 GTGAGCACGT TCTGCTCGCT CGCCAGGTCG GCGTTCCTTA CATCCTCGTT 100  
 GCTCTGAACA AGTGCGACAT GGTTGATGAT GAGGAAATCA TCGAGCTCGT 150  
 CGAGATGGAG ATCCRTGAGC TGCTCGCTGA GCAGGATTAC GACGAAGAGG 200  
 55 CTCCAATCAT CCACATCTCC GCACTGAAGG CTCTTGAGGG CGACGAGAAG 250  
 TGGACCCAGT CCATCATCGA CCTCATGCAG GCTTGCKATG ATTCCATCCC 300  
 AGACCCAGAG CGTGAGACCG ACAAGCCATT CCTCATGCCT ATCGAGGACA 350  
 TCTTCACCAT CACCGGCCGC GGTACCGTTG TTACCGGCCG TGTGAGCGT 400  
 GGCTCCCTGA AGGTCAACGA GGACGTCGAG ATCATCGGTA TCCGCGAGAA 450  
 60 KGCTACCACC ACCACCGTTA CCGGTATCGA GATGTTCCGT AAGCTTCTCG 500

```

ACTACACCGA GGCTGGCGAC AACTGTGGTC TGCTTCTCCG TGGCGTTAAG 550
CGCGAAGACG TTGAGCGTGG CCAGGTGTGT GTTAAGCCAG GCGCTTACAC 600
CCCTCACACC GAGTTCGAGG GCTCTGTCTA CGTTCGTGTCC AAGGACGAGG 650
GTGGCCGCCA CACCCCATTC TTCGACAACT ACCGCCCACA GTTCTACTTC 700
5 CGCACCACCG ACGTTACCGG TGTTGTGAAG CTTCTGAGG GCACCGAGAT 750
GGTCATGCCT GCGGACAACG TCGACATGTC CGTCACCCTG ATCCAGCCTG 800
TCGCTATGGA TGAGGGCCTG CGCTTCGCTA TC 832

```

10

## 2) INFORMATION FOR SEQ ID NO: 663

## (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 1192 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida catenulata*
- (B) STRAIN: ATCC 10565

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 663

```

AACGGCGAGC AAGACTTGGT GTTGGAGGTT TCTCAGCACT TGGGTGAGAA 50
CACCGTGCGT ACCATTGCCA TGGACGGTAC CGAGGGTTTG GTGAGAGGTA 100
CCGCTGTCAC TGACACCGGT GCTCCCATCT CGGTCCCCGT TGGTCAGGGT 150
30 ACCTTGGGCC GGATCATCAA CGTTGTCCGT GAGCCCATCG ACGAGCGTGG 200
TCCCATCCAG TGCAAGCAGA GAAACCCCAT TCACGCCGAG CCCCCGTCTT 250
TCACCGAGCA GTCCGTCGAG GCTGAGGTGT TGGAGACCGG TATCAAGGTT 300
GTCGACTTGT TGGCTCCCTA CGCCCGTGGT GGTAAGATTG GTCTTTTCGG 350
TGGTGCCGGT GTCGGTAAGA CCGTGTTTAT CCAGGAGTTG ATTAACAACA 400
35 TTGCCAAGGC CCACGGTGGT TTCTCCGTGT TCACTGGTGT CCGTGAGCGT 450
ACTCGTGAGG GTAACGACTT GTACCGTGAA ATGAAGGAGA CCGGTGTCAT 500
CAACTTGAGG GGCAGACTCA AGGTGGCCTT GGTGTTCCGT CAGATGAACG 550
AGCCCCCGGG GGCTCGTGCC CGTGTCGCCT TGACCGGTTT GACCATTGCC 600
GAGTACTTCC GTGACGAGGA AGGCCAGGAC GTGTTGTTGT TCGTTGACAA 650
40 CATTTTCAGA TTCACCCAGG CCGGTTCCGA GGTGTCGGCG TTGTTGGGTC 700
GTATCCCCTC CGCCGTCGGT TACCAGCCCA CTTTGGCCAC CGACATGGGT 750
TTGTTGCAGG AGAGAATTAC CACCACCAAG AAGGGTTCCG TCACCTCTGT 800
GCAGGCCGTG TACGTCCCTG CCGATGACTT GACTGACCCT GCCCCGCCA 850
CCACTTTCGC TCACTTGGAC GCCACCACCG TGTTGTCGCG TGGTATCTCC 900
45 GAGTTGGGTA TCTACCCCGC CGTCGACCCC TTGGACTCCA AGTCGAGATT 950
GTTGGACGTC GAGGTTGTTG GCCAGGAGCA CTACGACGTC GCCACCGGTG 1000
TCCAGGAGTG CTTGCAGGCC TACAAGTCGT TGCAGGACAT CATTGCCATT 1050
TTGGGTATGG ACGAGTTGTC CGAGCAGGAC AAGTTGACCG TCGAGAGAGC 1100
CCGTAAGATC CAGCGTTTCT TGTCGCAGCC CTTGCTGTC GCCGAGGTTT 1150
50 TCACTGGTAT CCCCGGTAGA TTGGTGAGAT TGCAGGACAC CG 1192

```

55

## 2) INFORMATION FOR SEQ ID NO: 664

## (i) SEQUENCE CHARACTERISTICS:

60

- (A) LENGTH: 29 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear



(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 664

5

AAYATGATIA CIGGIGCIGC ICARATGGA

29

10 2) INFORMATION FOR SEQ ID NO: 665

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1377 bases

(B) TYPE: Nucleic acid

15 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(C) ACCESSION NUMBER: extracted from X00779

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 665

25 ATGGGTAAAG AGAAGTCTCA CATTAAACGTT GTCGTTATCG GTCATGTCGA 50  
 TTCTGGTAAG TCTACCACTA CCGGTCATTT GATTTACAAG TGTGGTGGTA 100  
 TTGACAAGAG AACCATCGAA AAGTTCGAAA AGGAAGCCGC TGAATTAGGT 150  
 AAGGGTTCTT TCAAGTACGC TTGGGTTTTG GACAAGTTAA AGGCTGAAAG 200  
 30 AGAAAGAGGT ATCACTATCG ATATTGCTTT GTGGAAGTTC GAAACTCCAA 250  
 AGTACEAAGT TACCGTTATT GATGCTCCAG GTCACAGAGA TTTCATCAAG 300  
 AACATGATTA CTGGTACTTC TCAAGCTGAC TGTGCTATCT TGATTATTGC 350  
 TGGTGGTGTG GGTGAATTCG AAGCCGGTAT CTCTAAGGAT GGTCAAACCA 400  
 GAGAACACGC TTTGTTGGCT TTCACCTTGG GTGTTAGACA ATTGATTGTT 450  
 35 GCTGTCAACA AGATGGACTC CGTCAAAATGG GACGAATCCA GATTCCAAGA 500  
 AATTGTCAAG GAAACCTCCA ACTTTATCAA GAAGGTTGGT TACAACCCAA 550  
 AGACTGTTCC ATTCGTCCCA ATCTCTGGTT GGAACGGTGA CAACATGATT 600  
 GAAGCTACCA CCAACGCTCC ATGGTACAAG GGTGGGAAA AGGAAACCAA 650  
 GGCCGGTGTG GTCAAGGGTA AGACTTTGTT GGAAGCCATT GACGCCATTG 700  
 40 AACAACCATC TAGACCAACT GACAAGCCAT TGAGATTGCC ATTGCAAGAT 750  
 GTTTACAAGA TTGGTGGTAT TGGTACTGTG CCAGTCGGTA GAGTTGAAAC 800  
 CCGGTGTCATC AAGCCAGGTA TGGTTGTTAC TTTTGCCCCA GCTGGTGTGA 850  
 CCACTGAAGT CAAGTCCGTT GAAATGCATC ACGAACAATT GGAACAAGGT 900  
 GTTCCAGGTG ACAACGTTGG TTTCAACGTC AAGAACGTTT CCGTTAAGGA 950  
 45 AATCAGAAGA GGTAACGTCT GTGGTGACGC TAAGAACGAT CCACCAAAGG 1000  
 GTTGCGCTTC TTTCAACGCT ACCGTCATTG TTTTGAACCA TCCAGGTCAA 1050  
 ATCTCTGCTG GTTACTCTCC AGTTTTGGAT TGTCACACTG CTCACATTGC 1100  
 TTGTAGATTG GACGAATTGT TGGAAAAGAA CGACAGAAGA TCTGGTAAGA 1150  
 AGTTGGAAGA CCATCCAAAG TTCTTGAAGT CCGGTGACGC TGCTTTGGTC 1200  
 50 AAGTTCGTTT CATCTAAGCC AATGTGTGTT GAAGCTTTCA GTGAATACCC 1250  
 ACCATTAGGT AGATTCGCTG TCAGAGACAT GAGACAAACT GTCGCTGTCTG 1300  
 GTGTTATCAA GTCTGTTGAC AAGACTGAAA AGGCCGCTAA GGTTACCAAG 1350  
 GCTGCTCAAA AGGCTGCTAA GAAATAA 1377

55

2) INFORMATION FOR SEQ ID NO: 666

(i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 1536 bases

578

(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*  
 (C) ACCESSION NUMBER: extracted from M12082

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 666

	ATGGTTTTGC	CAAGACTATA	TACTGCTACA	TCCCGTGCTG	CTTTTAAAGC	50
	AGCCAAACAA	TCCGCTCCGC	TTCTATCCAC	TTCGTGGA	AGATGTATGG	100
15	CCTCAGCTGC	TCAATCTACT	CCAATCACCG	GTAAAGTTAC	CGCTGTCATT	150
	GGTGCCATTG	TTGACGTTCA	TTTTGAACAA	TCAGAGTTGC	CCGCTATTTT	200
	GAACGCTTTA	GAAATTAAAA	CACCTCAAGG	TAAGTTGGTT	TTGGAAGTTG	250
	CTCAACATTT	GGGTGAAAAC	ACTGTCAGAA	CCATTGCTAT	GGATGGTACC	300
	GAAGGTTTGT	TCCGTGGTGA	AAAGGTTCTT	GACACTGGTG	GCCCTATCTC	350
20	CGTCCCAGTT	GGGAGAGAAA	CTTTAGGGAG	AATCATCAAC	GTTATCGGTG	400
	AACCTATTGA	TGAAAGAGGT	CCAATTAAGT	CCAAACTAAG	AAAGCCAATT	450
	CACGCAGACC	CTCCTAGTTT	TGCAGAACAA	TCTACTTCGG	CTGAAATTTT	500
	GGAACAGGT	ATCAAAGTCG	TCGATCTATT	AGCTCCTTAT	GCCAGAGGTG	550
	GTAAGATTGG	TCTTTTCGGT	GGTGCAGGTG	TCGGTAAGAC	TGTGTTTCATT	600
25	CAAGAATTGA	TTAACAATAT	CGCCAAGGCC	CATGGTGGTT	TTTCCGTTTT	650
	CGCCGGTGT	GGTGAAAGGA	CCAGAGAGGG	TAATGACTTG	TACCGTGAAA	700
	TGAAGGAAAC	TGGAGTCATT	AACTTGGAAG	GTGAATCCAA	GGTCGCCTTA	750
	GTTTTTCGGTC	AAATGAACGA	ACCTCCAGGA	GCCAGAGCCA	GAGTCGCTTT	800
	AACTGGTTTG	ACGATCGCTG	AATATTTTTCAG	AGATGAAGAA	GGTCAAGACG	850
30	TCTTGTTGTT	TATCGACAAT	ATCTTTTAGAT	TTACTCAAGC	TGGTTCAGAA	900
	GTCTGTGCGC	TTTTTGGGTCG	TATTCATCT	GCCGTCGGTT	ATCAACCAAC	950
	TTTGGCCACT	GATATGGGTC	TCTTACAAGA	AAGAATTACC	ACCACAAAGA	1000
	AGGGTTCTGT	CACCTCTGTG	CAAGCCGTTT	ATGTTCCAGC	CGATGATTTA	1050
	ACAGATCCGT	CTCCGTCCAC	ATCTTTTGCC	CATTTGGACG	CATCATCCGT	1100
35	CTTGTCAGA	GGTATTTTCAG	AATTAGGTAT	TTACCCTGCA	GTGGATCCAT	1150
	TGGATTCTAA	ATCAAGGTTA	TTGGATGCCG	CCGTTGTCGG	TCAAGAACAT	1200
	TATGACGTCG	CCTCCAAGGT	TCAAGAACT	TTACAGACCT	ATAAATCTTT	1250
	ACAAGATATC	ATTGCTATTT	TGGGTATGGA	TGAATTGTCC	GAACAAGATA	1300
	AACTAACTGT	CGAAAGGGCA	AGAAAGATTC	AAAGATTCTT	ATCTCAACCA	1350
40	TTTGCTGTGCG	CCGAAGCTCT	TACTGGTATC	CCAGGTAAAT	TAGTGAGATT	1400
	AAAGGACACC	GTTGCCTGTT	TCAAAGCCGT	TTTGGAAGGT	AAATACGATA	1450
	ATATACCAGA	ACATGCTTTC	TATATGGTTG	GTGGTATTGA	AGATGTTGTT	1500
	CGTAAAGCTG	AAAAGTTAGC	CCGTGAAGCC	AACTAG		1536

45

2) INFORMATION FOR SEQ ID NO: 667

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 1293 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

55 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma cruzi*  
 (B) STRAIN: Y

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 667

	CGGAAGGCGT	GCCGCCGGTG	TTAACTGCTT	TGGATGTGGT	GGAAAACTT	50
	GGCCGTGACG	AGCCCTTGAC	ACTTGAAATT	GTACAGCACT	TGGACGCCCCA	100
5	CACCGGCCGT	TGCATTGCCA	TGCAGACGAC	AGATCTGCTT	AAGCTGAAAT	150
	CGAAGGTAGT	CTCGAGTGGC	GGGAACATTT	CTGTGCCTGT	GGGTCTGGGAG	200
	ACACTTGGCA	GAATCTTCAA	TGTGCTGGGC	GATGCCATTG	ACCAGCGCGG	250
	CCATGTTGGA	GAGAAGCAAC	GCATGCCTAT	TCACGCTGTG	GCACCAAAGT	300
	TGGCGGATCA	GGCGGCAGAG	GACACCATTG	TCACAACGGG	TATCAAGGTG	350
10	ATTGATCTGA	TTTTGCCCTA	CTGCAAGGGT	GGGAAGATTG	GTCTTTTGG	400
	TGGTGTCTGC	GTGGGCAAGA	CGGTTATTAT	CATGGAACTT	ATTAACAACG	450
	TTGCCAAGGG	CCATGGTGGG	TTCTCCGTCT	TTGCGGGTGT	TGGTGAACGC	500
	ACGCGTGAGG	GGACGGATCT	TTACCTGGAG	ATGATGCAAT	CCAAAGTTAT	550
	TGACCTGAAG	GGTGACTCGA	AGTGTGTGTT	GGTGTATGGT	CAGATGAACG	600
15	AGCCTCCGGG	TGCCCGTGCT	CGTGTGGCGC	AGTCTGCCTT	GACAATGGCC	650
	GAGTACTTCC	GTGATGTGGA	AGGGCAAGAC	TGTGTGCTTT	TTATCGACAA	700
	CATTTTCCGC	TTTACGCAGG	CAAACTCTGA	GGTGTGAGCG	CTGTTGGGTC	750
	GTATTCCCGC	TGCCGTGCGC	TACCAGCCTA	CCCTTGCTGA	GGATCTTGGG	800
	CAGTTGCAGG	AGCGCATTAC	GTCCACGACA	AAAGGTTCCA	TTACCTCTGT	850
20	GCAGGCTGTG	TACGTGCCAG	CCGATGATAT	TACCGACCCT	GCGCCAGCAA	900
	CGACCTTTTC	CCACCTCGAT	GCCACGACGG	TGCTGGACCG	TGCCGTTGCC	950
	GAATCCGGCA	TTTACCCCGC	TGTCAACCCA	CTGGAGTGTG	CCTCGCGTAT	1000
	CATGGACCCG	GATGTAATCA	CGGTTGACCA	CTACAACGTG	GCGCAGGATG	1050
	TGGTGCAGAT	GCTTACCAAG	TACAAGGAGC	TGCAGGATAT	CATTGCGGTG	1100
25	CTTGGCATTG	ATGAGCTCAG	TGAGGAGGAT	AACTTATCG	TGGACCGTGC	1150
	GCGTAAGGTG	ACAAAGTTTC	TCTCCAGGCC	TTTCCAGGTG	GCGGAGGTGT	1200
	TTACTGGCAT	GACAGGCCAC	TACGTGCAGC	TGGAGGAGAC	CATTGAGTCT	1250
	TTTTCCGGCC	TGTTGATGGG	CACATATGAT	CAGGTGCCCG	AGA	1293

30

## 2) INFORMATION FOR SEQ ID NO: 668

## (i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 1191 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 40 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Corynebacterium glutamicum*  
 (C) ACCESSION NUMBER: X77034

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 668

	GTGGCAAAGG	CGAAGTTCGA	GCGTACCAAG	CCCCACGTAA	ACATCGGCAC	50
	CATCGGTCAC	GTTGACCACG	GTAAGACCAC	CACCACCGCG	GCTATCACCA	100
50	AGGTTCTGGC	TGACACTTAC	CCTGAGCTCA	ACGAGGCTTT	CGCCTTCGAC	150
	TCCATCGATA	AGGCTCCTGA	GGAGAAGGAG	CGTGGCATCA	CGATCAACAT	200
	CTCCCACGTT	GAGTACCAGA	CTGAAAAGCG	CCACTACGCA	CACGTTGACG	250
	CTCCAGGCCA	CGCCGACTAC	ATCAAGAACA	TGATTACCGG	CGCTGCTCAG	300
	ATGGACGGCG	CAATCCTCGT	TGTTGCTGCT	ACCGACGGCC	CAATGCCTCA	350
55	GACCCGTGAG	CACGTTCTTC	TTGCTCGCCA	GGTTGGCGTT	CCTTACATCC	400
	TCGTTGCTCT	TAACAAGTGC	GACATGGTTG	AGGATGAGGA	AATCATCGAG	450
	CTCGTCGAGA	TGGAAGTTTC	TGAACTTCTT	GCTGAGCAGG	ACTACGACGA	500
	AGAGGCTCCA	ATTGTTTACA	TCTCCGCTCT	GAAGGCTCTT	GAGGGCGACG	550
	AGAAGTGGGG	CAAGCAGATC	CTTGAGCTCA	TGCAGGCTTG	CGATGACAAC	600
60	ATCCCTGACC	CAGTTCGTGA	GACCGACAAG	CCATTCCTCA	TGCCTATCGA	650

```

GGACATCTTC ACCATCACCG GTCGTGGCAC CGTTGTTACC GGTCGTGTTG 700
AGCGCGGTAC CCTGAACGTG AACGATGATG TTGACATCAT CGGCATCAAG 750
GAGAAAGTCCA CCTCCACCAC CGTTACCGGT ATCGAGATGT TCCGTAAGCT 800
TCTTGACTCC GCTGAGGCTG GCGACAACCTG TGGTCTGCTT CTCCGTGGTA 850
5 TCAAGCGCGA AGATGTTGAG CGTGGCCAGG TTATCGTTAA GCCAGGCGCT 900
TACACCCCTC ACACCGAGTT CGAGGGCTCT GTCTACGTTT TGTCCAAGGA 950
TGAAGGTGGC CGCCACACCC CATTCCTCGA CAACTACCGT CCTCAGTTCT 1000
ACTTCCGCAC CACCGACGTT ACCGGTGTG TGAAGCTTCC AGAGGGCACC 1050
GAGATGGTCA TGCCTGGCGA CAACGTCGAC ATGTCCGTCA CCCTGATCCA 1100
10 GCCTGTCGCT ATGGACGAGG GCCTGCGTTT CGCTATCCGC GAAGGCTCCC 1150
GCACCGTTGG CGCTGGTCGT GTCACCAAGA TCATCAAGTA A 1191

```

15 2) INFORMATION FOR SEQ ID NO: 669

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 1383 bases
(B) TYPE: Nucleic acid
20 (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

25 (vi) ORIGINAL SOURCE:

```

(A) ORGANISM: Escherichia coli
(C) ACCESSION NUMBER: extracted from V00267

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 669

```

30 ATGGCTACTG GAAAGATTGT CCAGGTAATC GGCGCCGTAG TTGACGTCTGA 50
ATTCCCTCAG GATGCCGTAC CGCGCGTGTA CGATGCTCTT GAGGTGCAAA 100
ATGGTAATGA GCGTCTGGTG CTGGAAGTTC AGCAGCAGCT CGGCGGCGGT 150
ATCGTACGTA CCATCGCAAT GGGTTCCTCC GACGGTCTGC GTCGCGGTCT 200
35 GGATGTAAAA GACCTCGAAC ACCCGATTGA AGTCCCGGTA GGTAAGCGA 250
CTCTGGGCCG TATCATGAAC GTACTGGGTG AACCGGTCGA CATGAAAGGC 300
GAGATCGGTG AAGAAGAGCG TTGGGCGATT CACCGCGCAG CACCTTCCTA 350
CGAAGAGCTG TCAAACCTCTC AGGAACTGCT GGAAACCGGT ATCAAAGTTA 400
TCGACCTGAT GTGTCCGTTT GCTAAGGGCG GTAAAGTTGG TCTGTTCGGT 450
40 GGTGCGGGTG TAGGTAAAAAC CGTAAACATG ATGGAGCTCA TTCGTAACAT 500
CGCGATCGAG CACTCCGGTT ACTCTGTGTT TGCGGGCGTA GGTGAACGTA 550
CTCGTGAGGG TAACGACTTC TACCACGAAA TGACCGACTC CAACGTTATC 600
GACAAAGTAT CCCTGGTGTA TGGCCAGATG AACGAGCCGC CGGGAAACCG 650
TCTGCGCGTT GCTCTGACCG GTCTGACCAT GGCTGAGAAA TTCCGTGACG 700
45 AAGGTCGTGA CGTTCTGCTG TTCGTTGACA ACATCTATCG TTACACCCTG 750
GCCGGTACGG AAGTATCCGC ACTGCTGGGC CGTATGCCTT CAGCGGTAGG 800
TTATCAGCCG ACCCTGGCGG AAGAGATGGG CGTTCTGCAG GAACGTATCA 850
CCTCCACCAA AACTGGTTCT ATCACCTCCG TACAGGCAGT ATACGTACCT 900
CGGGATGACT TGA CTGACCC GTCTCCGGCA ACCACCTTTG CGCACCTTGA 950
50 CGCAACCGTG GTACTGAGCC GTCAGATCGC GTCTCTGGGT ATCTACCCGG 1000
CCGTTGACCC GCTGGACTCC ACCAGCCGTC AGCTGGACCC GCTGGTGGTT 1050
GGTCAGGAAC ACTACGACAC CGCGCGTGGC GTTCAGTCCA TCCTGCAACG 1100
TTATCAGGAA CTGAAAGACA TCATCGCCAT CCTGGGTATG GATGAACTGT 1150
CTGAAGAAGA CAAACTGGTG GTAGCGCGTG CTCGTAAGAT CCAGCGCTTC 1200
55 CTGTCCACAG CGTCTTCGT GGCAGAAGTA TTCACCGGTT CTCCGGGTAA 1250
ATACGTCTCC CTGAAAGACA CCATCCGTGG CTTTAAAGGC ATCATGGAAG 1300
GCGAATACGA TCACCTGCCG GAGCAGGCGT TCTACATGGT CGGTTCCATC 1350
GAAGAAGCTG TGGAAAAAGC CAAAAAACTT TAA 1383

```

60

## 2) INFORMATION FOR SEQ ID NO: 670

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1410 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*  
 (B) STRAIN: NCTC 11638  
 (C) ACCESSION NUMBER: extracted from AF004014

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 670

20	ATGAAAGCGA	TGGAAGGTAA	AATCATTCAG	GTTTTAGGCC	CGGTGGTAGA	50
	TGTGGAGTTT	GAATCCTATC	TGCCGGCGAT	TTTTGAAGCA	CTAGACATTA	100
	ATTTTGAAGT	TAATGGCGTT	CAAAAATCTT	TAGTTTTAGA	GGTGGCAGCC	150
	CATTTGGGTG	GTAATCGGGT	GCGAGCGATT	GCTATGGATA	TGACAGAAGG	200
	CTTAGTGCCT	AACCAAGCCG	TCAAAGCTCG	TGGCAAATG	ATTGAAGTGC	250
	CTGTGGGCGA	AGAAGTGTTA	GGGCGTATTT	TTAATGTTGT	GGGCGAGAGC	300
25	ATTGATAATT	TAGAGCCGCT	TAAGCCGTCC	TTAACTTGGC	CCATTCACAG	350
	AAAAGCCCCCT	AGTTTTGAGC	AGCAAAGCAC	TAAAACAGAA	ATGTTTGAAA	400
	CCGGTATTAA	AGTCATTGAC	TTGCTCGCGC	CTTATTCTAA	GGGCGGTAAA	450
	GTAGGCTTGT	TTGGTGGGGC	TGGCGTAGGC	AAAACGGTGA	TCATTATGGA	500
	GCTTATCCAC	AATGTGGCTT	ATAAGCATAA	CGGGTATTCT	GTGTTTGCAG	550
30	GTGTGGGGGA	GCGCACCAGA	GAAGGGAACG	ATCTGTATTT	TGAGATGAAA	600
	GAAGGGGGCG	TTTTAGACAA	AGTTGCGTTG	TGCTATGGGC	AAATGAATGA	650
	GCCACCAGGT	GCAAGGAATC	GCATCGCATT	CACCGGCTTG	ACGATGGCGG	700
	AGTATTTCCG	TGATGAAAAG	GGCTTAGATG	TGTTGATGTT	TATTGATAAC	750
	ATCTTTAGAT	ACGCTCAAAG	CGGTGCGGAA	ATGAGCGCGC	TATTAGGCCG	800
35	TATCCCTTCA	GCGGTGGGGT	ATCAGCCCA	GCTAGCCGGG	GAAATGGGGA	850
	AACTTCAAGA	CGGTATCGCT	TCCACTAAAA	ATGGCTCTAT	CACTTCGGTT	900
	CAAGCGGTGT	ATGTGCCAGC	AGACGACTTG	ACTGACCCAG	CCCCTGCTTC	950
	GGTGTTTGCG	CATTTAGATG	CGACTACGGT	GTTGAATAGA	AAGATCGCTG	1000
	AAAAAGGGAT	TTATCCGGCG	GTGGATCCTT	TGGATTCCAC	TTCAAGGATT	1050
40	TTAAGCCCTC	AAATGATTGG	CGAGAAGCAC	TATGAAATCG	CCACCGGTAT	1100
	CCAGCAAGTT	TTGCAAAAAT	ACAAGGATTT	GCAAGATATT	ATTGCGATTT	1150
	TGGGATTGGA	CGAATTGAGC	GAAGAGGATC	AAAAAACGGT	TGAAAGGGCC	1200
	AGAAAAATTG	AGAAGTTTTT	ATCCCAGCCG	TTTTTTGTGG	CTGAAGTGTT	1250
	TACAGGAAGT	CCCGGTAAGT	ATGTGACTCT	CCAAGAGACT	TTAGAGGGCT	1300
45	TTGGAGGGAT	TTTAGAGGGC	AAATACGATC	ACATTCCTGA	AAACGCGTTT	1350
	TACATGGTGG	GCAGCATTCA	AGAGGTTTTA	GAAAAAGCTA	AAAACATGAA	1400
	AAATTCCTAA					1410

## 2) INFORMATION FOR SEQ ID NO: 671

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1401 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium acetobutylicum*  
 (B) STRAIN: DSM 792  
 (C) ACCESSION NUMBER: extracted from AF101055

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 671

```

10  ATGCCAGAAC ATGTAGGTAA AATTGTTTCAG GTAATAGGAC CTGTTGTGGA      50
    TATTAAATTT GATGCAGAGA ACCTTCCTGA CATCTATAAT TCCATAGAAA      100
    TAGATATGGG AGATAATAAA AAACCTCATTG CTGAAGTTGA ACAACATGTA      150
    GGAGATGACA TAGTAAGAAC AATAGCAATG GAAGGTACTG ACGGATTAAA      200
    AAGAGGAATG GAAGCAGTTA ACACTGGTAA ACCAATATCT GTACCAGTTG      250
    GAGAAAATGT TTTAGGACGT CTTTTTAATG TTTTAGGTCA GACAATAGAT      300
    GAAGCAGGAG ACATGAATGC TGATAAGTAT TATCCAATTG ATAGACCAGC      350
15  TCCAACCTTT GAAGAACAAT CAGTTCAACC AGAAATGTTT GAAACAGGTA      400
    TTAAGGTTAT AGATTTACTT GCTCCATATC AAAAGGGTGG AAAAATCGGT      450
    TTGTTTCGGT GTGCCGGTGT TGGTAAAACA GTTCTTATTC AGGAACTTAT      500
    AAATAATATA GCAAAAGAAC ACGGTGGATT ATCAGTATTC ACAGGTGTTG      550
    GAGAAAGAAC AAGAGAAGGT AATGACCTTT ATTATGAAAT GAAAGATTCA      600
20  GGAGTTATAA ATAAACAGC TCTAGTATTT GGTCAGATGA ATGAACCACC      650
    TGGCGCTAGA ATGAGAGTTG CTTTAACAGG ACTTACAATG GCTGAATATT      700
    TTAGAGACAA AGGTCAAGAT GTGCTTCTAT TTATAGATAA TATATTCAGA      750
    TTTACACAAG CTGGTTCAGA GGTTTCAGCT TTAAGTGGTA GAATACCTAG      800
    TGCCGTTGGT TATCAGCCAA CACTTGCAAA TGAAATGGGT GCTCTTCAAG      850
25  AGAGAATAAC ATCAACAAAA CAGGGTTCAA TCACATCCGT TCAGGCTGTA      900
    TATGTTCCCT CTGATGACCT TACAGACCCA GCTCCAGCAA CAACATTTAC      950
    GCATCTTGAT GCAACAACAG TTCTTTCAAG AGAAATATCA AACTTAGGAA      1000
    TATATCCTGC AGTTAGTCCT CTTGAATCAA CTTCAAGAAT ACTTGATCCA      1050
    AGAATTGTTG GAGAAGAGCA TTATGAAGTT GCTAACAAGG TTAAACATAT      1100
30  ACTTGAAAGA TATCAAGAAC TTCAAGATAT CATAGCTATA CTTGGTGTTG      1150
    ATGAACTTTC AGATGAGGAT AGATTGTTAG TTGGAAGAGC AAGAAGAGTA      1200
    CAGAGATTCT TATCTCAAGC TTTTAGTGTT GCTGAACAAT TTACAGGAAT      1250
    GAAAGGTCAG TTTGTACCTG TAAAAGATAC TATAAGAAGT TTTAAAGAAA      1300
    TATTAGATGG TAAGTGTGAT GATCTTCCAG AAGCTGCATT TTTATTTGCA      1350
35  GGAACAATAG AAGATGTAAA AGAAAAAGCT AAAAAAATGA TGGAAAGCTA      1400
    A
  
```

## 40 2) INFORMATION FOR SEQ ID NO: 672

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1509 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 50 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cytophaga lytica*  
 (B) STRAIN: DSM 2039  
 (C) ACCESSION NUMBER: M22535

## 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 672

```

60  ATGTCTAAAG TTACAGGTAA AGTTTCCCAA ATTATTGGCC CAGTTATAGA      50
    TGTGGAGTTC CAAGCAGGGG TAGATCTTCC AAAAATTTAT GATTCATTAG      100
    AAATTAAAAA AGCAGATGGA TCAATTTTGG TTTTGGAAGT ACAATCACAC      150
    ATTGGTGAGA ACACAGTAAG AACTATATCT ATGGATTCTT CTGATGGTTT      200
  
```

```

AAGTCGTGGA GCAGAGGTTA ATGCAACAGG AAGCGCTATA CAAATGCCAG 250
TTGGAGATGA CGTTTACGGA CGTTTATTTA ACGTAATTGG AGACGCTATT 300
GATGGTCTTG GGAATTTACC AAAATCTGGT AAAGATGGTT TGCCAATACA 350
CAGAGAGGCA CCAAAATTTG AAGACTTATC TACTTCTACA GAAGTATTAT 400
5 TTACAGGTAT TAAAGTAATT GACCTTATTG AGCCTTATGC AAAAGGTGGT 450
AAGATTGGTT TATTTGGAGG TGCCGGAGTA GGTAACACAG TATTAATTCA 500
GGAATTAATT AACAACATTG CAAAAGGTCA CCGTGGACTT TCTGTATTTG 550
CTGGTGTAGG TGAGCGTACT CGTGAGGGTA ACGATTTACT ACGTGAGATG 600
TTAGAGTCTG GTATTATTAA ATACGGAGAT GACTTTATGC ACTCTATGGA 650
10 AGAAGGTGGT TGGGATTTAT CTAAAGTTGA TAAATCTGTA ATGAAAGATT 700
CTAAAGCAAC CTTTGTATTT GGACAAATGA ATGAGCCACC AGGAGCACGT 750
GCACGTGTTG CATTATCTGG TTAACTATT GCAGAATATT TCCGTGATGG 800
AGCAGGTGAA GGTCAGGTA AAGATGTACT TTTCTTTGTG GATAACATTT 850
TCCGTTTTAC ACAAGCTGGT TCTGAGGTAT CTGCATTACT TGGTCGTATG 900
15 CCATCTGCGG TAGGTTACCA ACCAACATTA GCAACAGAAA TGGGTGCTAT 950
GCAAGAGCGT ATTACATCAA CAAAAGAGG TTCTATTACA TCTGTACAGG 1000
CGGTTTACGT ACCTGCGGAT GATTTAACGG ATCCAGCACC GGCAACTACC 1050
TTTGCTCACT TAGATGCAAC AACGGTATTG TCTCGTAAAA TTGCAGAGTT 1100
AGGTATTTAC CCAGCGGTAG ATCCATTAGA TTCTACTTCT AGAATCTTAG 1150
20 CTCCAGAAAT TTTAGGAAAA GATCACTACT CTTGTGCACA ACGTGTAATA 1200
GAGTTGTTAC AACGTTATAA AGAATTACAA GATATTATTG CTATCCTTGG 1250
TATGGAAGAA TTATCTGAGG AAGATAAAAT GGCAGTTGGT AGAGCAAGAC 1300
GTGTACAACG TTTCTTATCT CAGCCTTTCC ACGTAGCAGA GCAATTTACA 1350
GGTCTTAAAG GTGTTTTAGT AGATATCAAG GATACTATTA AAGGATTTAA 1400
25 TATGATTATG GATGGTGAAT TAGATCACTT ACCAGAATCT GCATTTAACC 1450
TTAAAGGTAC TATTGAAGAA GCTATAGAAG CAGGAGAAAA AATGCTTGCT 1500
GAAGCATAA 1509

```

30

## 2) INFORMATION FOR SEQ ID NO: 673

## (i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 819 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

40

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Ehrlichia risticii*
- (B) STRAIN: HRC-IL

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 673

```

CCTAAAATAT ATGATGCATT ATATGTAAAA CTAGATAATG AAAATTTGTG 50
TTTAGAAGTT TCACAAATTA TTGGAGATAA TGTTGTTAGA TGTATTGCAA 100
TGGGAGCTAC TTATGGATTA AATCGTGGTT TAGAAGTAGT TTGTTTCAGGA 150
50 AATCCAATTC AGGTTCCCTGT AGGTGAACAA GTTTTAGGTA GAATGTTTAA 200
TGTTGTTGGT AAAACAATTG ACAATCTTGA ATCTTTAGAT GATAAAAATA 250
TAAAAATGAT GCCAATTCAT CGAAATCCAC CATCATTTGA AGAGCAATCC 300
AATGAAATG AAATTTTGA AACAGGCATT AAAGTTATTG ATTTATTAAT 350
TCCATATGCT AAAGGTGGTA AGATTGGATT ATTTGGTGGA GCAGGGGTTG 400
55 GGAAAACGGT TCTTGTTCAA GAATTAATTC ACAATATCGC AAAAGGTCAT 450
GGTGGTCTAT CTGTTTTTGC TGGAGTTGGT GAAAGAATC GTGAAGGTAA 500
TGACTTGTAT TATGAAATGA TTGAAGGTGG AGTTATAGAT AAAACAGCCT 550
TAGTGTTTGG GCAAATGAAT GAACCTCCTG GCGCAAGAAT GCGCGTAGCA 600
TTAACTGCTT TAACAATGGC TGAATATTTT CGTGATGTTT AAAACCAAGA 650
60 TGTTTTGTGA TTTATTGATA ATATCTTTAG ATTTACACAA GCTGGTAGTG 700

```

AAGTTTCAGC	ATTATTAGGA	AGAATGCCAT	CTGCTGTTGG	TTATCAACCA	750
ACTTTGGCAT	ATGAAATGGG	ATTGTTACAA	GAAAGAATCA	CTTCCACTAA	800
AAGTGGTTCT	ATAACATCT				819

5

## 2) INFORMATION FOR SEQ ID NO: 674

## (i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 840 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Vibrio cholerae*  
 (B) STRAIN: ATCC 25870

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 674

AGAGCGAAGT	ACCAAGTGTT	TACGATGCTC	TGAATGTTGT	GGATTCCAAA	50
GAACGTCTGG	TTCTGGAAGT	TCAACAGCAG	CTAGGCGGTG	GCGTGATTCTG	100
25 CGCTATCGTT	ATGGGTAGCT	CGGATGGTTT	ACGTCGTGGA	ATGACAGTAC	150
AAAACACTGG	CGCTCCAATT	TCAGTACCAG	TAGGTACTAA	AACCCTAGGT	200
CGTATCATGA	ACGTGCTTGG	TGATGCGATT	GACGAACGTG	GCGACATTGG	250
CGCAGAAGAG	GTGTACTCGA	TTCACCGTCC	TGCTCCAAGC	TACGAAGAAC	300
AGTCTAGTGC	AAC TGAAC TT	TTGGAAACGG	GTGTTAAGGT	TATCGACCTG	350
30 ATCTGTCCGT	TTGCGAAAGG	CGGTAAAATC	GGTCTGTTCTG	GTGGTGCGGG	400
TGTAGGTAAG	ACCGTTAACA	TGATGGAAC T	TATCAACAAC	ATCGCGCTAC	450
AGCACTCAGG	TTTGTCAGTA	TTTGCTGGGG	TAGGTGAGCG	TACTCGTGAG	500
GGTAACGACT	TCTACCACGA	AATGCAGGAA	GCGGGCGTTG	TAAACGTTGA	550
ACAACCAGAA	CTGTCGAAAG	TAGCGATGGT	TTACGGTCAG	ATGAACGAGC	600
35 CACCAGGCAA	CCGTCTGCGT	G TAGCACTGA	CTGGTCTGAC	TATGGCGGAA	650
AAGTTCCGTG	ATGAAGGCCG	TGACGTACTG	CTGTTTATCG	ACAACATCTA	700
CCGTTACACC	CTAGCGGGAA	CGGAAGTATC	TGCTCTGCTT	GGCCGTATGC	750
CTTCAGCGGT	AGGTTACCAA	CCAACACTGG	CTGAAGAGAT	GGGTGTTCTG	800
40 CAAGAACGTA	TCACGTCAAC	CAAAAAAGGT	TCTATCACCT		840

## 2) INFORMATION FOR SEQ ID NO: 675

## (i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 828 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Vibrio cholerae*  
 (B) STRAIN: ATCC 25870

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 675

CGGCGGTATC	CTAGTTGTAG	CGGCAACTGA	CGGTCCAATG	CCACAAACTC	50
60 GTGAGCACAT	CCTGCTGGGT	CGCCAAGTAG	GTATTCCTTA	CATCATCGTA	100



	TTCATGAACA	AGTGTGACAT	GGTTGACGAT	GAAGAGCTTC	TAGAGCTGGT	150
	AGAGATGGAA	GTTCGTGAGC	TGCTGTCTGA	GTACGATTTT	CCAGGTGATG	200
	ACCTGCCAGT	AATCCAAGGT	TCAGCACTAG	GCGCGCTAAA	CGGCGAAGCA	250
	CAGTGGGAAG	CGAAGATTGT	TGAGCTAGCA	GAAGCACTGG	ATACTTAYAT	300
5	TCCAGAGCCA	GAGCGTGCAG	TAGACATGGC	ATTCTTGATG	CCAATCGAAG	350
	ACGTATTCTC	AATCCAAGGT	CGTGGTACAG	TAGTAAGTGG	CCGTATCGAG	400
	CGCGGCATCC	TGAAAGTGGG	TGACGAAGTA	GCGATCGTTG	GTATCAAAGA	450
	GACAGTAAAA	ACGACCTGTA	CAGGTGTAGA	GATGTTCCGT	AAGCTGCTTG	500
	ACGAAGGTCG	TGCAGGTGAG	AACGTAGGTG	CACTRCTACG	TGGTACTAAG	550
10	CGTGAAGAAG	TAGAGCGTGG	TCAAGTACTG	GCGAAGCCAG	GTTCAATCAC	600
	ACCACACACT	AAGTTCGAAT	CAGAAGTATA	CGTACTGTCA	AAAGATGAAG	650
	GTGGCCGTC	TACTCCATTC	TTCAAAGGTT	ACCGTCCACA	GTTCTACTTC	700
	CGTACAAC	ACGTAACAGG	CAGCATTGAG	CTACCAGARG	GCGTAGAAAT	750
	GGTAATGCCA	GGCGACAACG	TGAAGATGGT	TGTAGACCTG	ATTGCACCAA	800
15	TCGCGATGGA	CGAAGGTCTA	CGCTTCGC			828

## 2) INFORMATION FOR SEQ ID NO: 676

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1298 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania enriettii*  
 (B) STRAIN: ATCC 50120

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 676

35	CTTCTCGGAT	GGCGTGCCGC	CGGTGCTGAC	GGCGCTGGAC	GTGACGGAGG	50
	AACTCGGGCG	CGACGAGCCG	CTGACGCTAG	AGATCGTGCA	ACACTTGGAC	100
	GCGCACACCG	GCCGCTGCAT	TGCGATGCAG	ACGACGGACC	TACTGAAGCT	150
	GAAGTCGAAG	GTCGTGTCGA	CCGGCGGGAA	CATCTCTGTG	CCTGTGGGCC	200
	GCGAGACGCT	TGGCCGCATC	TTCAACGTAC	TGGGCGACGC	GATTGACCAG	250
40	CGCGGCCCCG	TGGGCGAGAA	GATGCGCATG	GCGATCCACG	CGGAGGCGCC	300
	GAAGCTGGCG	GACCAGGCGG	CGGAGGACAC	GATCCTGACG	ACCGGCATCA	350
	AAGTGATCGA	TCTGATCCTG	CCGTACTGCA	AGGGCGGCAA	GATCGGTCTG	400
	TTCCGGCGGTG	CCGGTGTGGG	GAAGACTGTG	ATCATCATGG	AGCTGATCAA	450
	CAACGTCGCG	AAGGGCCACG	GTGGCTTCTC	CGTGTTCCGC	GGCGTTGGGG	500
45	AGCGCACCCG	CGAGGGGACG	GATCTGTACC	TGGAGATGAT	GCAATCGAAG	550
	GTGATCGACC	TGAAAGGTGA	GTCGAAGTGT	GTGCTTGTGT	ACGGGCAGAT	600
	GAACGAGCCC	CCGGGTGCGC	GCGCGCGCGT	TGCGCAGTCT	GCGCTGACGA	650
	TGGCCGAGTA	CTTCCGCGAC	GTGGAGGGCC	AGAACGTGCT	GCTGTTTCATC	700
	GACAACATCT	TCCGCTTCAC	GCAGGCGAAC	TCCGAGGTGT	CAGCGCTGCT	750
50	GGGCCGCGATC	CCCGCCGCTG	TGGGCTACCA	ACCGACGCTT	GCGGAGGATC	800
	TCGGCATGTT	GCAGGAGCGC	ATCACGTCGA	CGACGAAGGG	GTCGATCACG	850
	TCTGTGCAGG	CGGTGTACGT	GCCTGCGGAT	GATATCACGG	ATCCC GCGCC	900
	GGCGACGACA	TTCTCGCACC	TGGACGCGAC	GACGGTGCTG	GACCGCGCGG	950
	TGGCGGAGTC	TGGCATCTAC	CCCGCGGTCA	ACCCACTGGA	GTGCGCGTCG	1000
55	CGTATCATGG	ACCCTGATGT	GATCGACGTG	GATCACTACA	ACGTTGCGCA	1050
	GGACATTGTT	CAGATGCTGA	CCAAGTACAA	GGAGCTGCAG	GATATCATTG	1100
	CGGTGCTTGG	TATCGACGAG	CTGAGCGAGG	AGGACAAGGT	TGTGGTGGAC	1150
	CGTGCGCGCA	AGGTGACCCG	GTTCTGTGTA	CAGCCGTTCC	AGGTTGCGGA	1200
	GGTGTTTACT	GGCATGACAG	GCCATTACGT	ACAGCTGAGC	GACACGGTGG	1250
60	AGTCGTTTCTC	CGGTTTGCTG	ATGGGGTCGT	ACGACCAGAT	TCCGGAGA	1298

## 2) INFORMATION FOR SEQ ID NO: 677

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1083 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia microtti*  
 (B) STRAIN: Persing-1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 677

20	CAAGCTCAAG TCTGAGCGTG AGAGAGGTAT TACTATTGAC ATTACTCTCT	50
	GGAAATTGGA GACCCAGAAA TACGAGTACA CTGTCATAGA CGCACCTGGT	100
	CATCGTGACT TTATCAAAAA TATGATTACT GGGACTTCAC AAGCCGACGT	150
	TGCTATGCTT GTCGTTCTTG CTGAATCTGG CGGATTCGAG GCTGCTTTTT	200
	CCAAAGAAGG TCAGACCCGT GAACACGCCT TACTAGCCTT CACACTTGGC	250
25	GTCAAACAGA TGATTGTTGC TATTAACAAA ATGGATTCTT GTCAGTACAA	300
	GGAGGATCGT TATATGGAAA TTTTCAAGGA AGTACAGCAG TACTTGAAGA	350
	AGGTGGGTGA CAAAGTTGAA AGCGTGCCGT TTGTTGCTAT TTCAGGATTC	400
	CACGGTGACA ACATGGTTGA AAAATCTACT AACATGCCTT GGTATAAGGG	450
	TAAGACCCTC GTAGAGGCAC TTGATCAAAT GGAGCCTCCA AAACGTCCGG	500
30	TCGAAAAACC TCTTAGATTG CCCCTGCAGT CAGTCTATAA AATTGGAGGT	550
	ATTGGTACGG TACCAGTCGG AAGGGTCGAA ACAGGACAAC TGAAAGCAGG	600
	AATGATCATT ACTTTTGCCC CCACTGGTTT GACCACTGAA TGTAATCTG	650
	TTGAAATGCA TCACGAGGTT GTGGAAGTGG CTAGCCCCGG TGATAACGTT	700
	GGATTTAATG TCAAGAATGT GTCTGTAAAG GATATTAAGA GAGGAAATGT	750
35	GGCTTCGGAT TCGAAAAATG ACCCAGCCAA GGAAGCTACC TCTTTCTCTG	800
	CACAAGTCAT GTTACTCAAT CACCCTGGTA CCATCAAGGC CGGTTACTCA	850
	CCTGTGGTTG ATTGCCATAC TGCCCACATT GCTTGCAAAT TCGAATCTCT	900
	AGACACTAGG ATTGACAAGC GTACTGGCAA GACTTTGGAA GAAAATCCTA	950
	AGACTATTAA GAATGGTGAC GCTGCCATGG TGACTATGAA ACCAAATAAA	1000
40	CCCATGGTTG TGGAAACTTT CACCGACTAC GCCCCGTTGG GCCGGTTCGC	1050
	CGTGCGTGAT ATGCGCCAAA CCGTTGCCGT CGG	1083

## 2) INFORMATION FOR SEQ ID NO: 678

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptococcus neoformans*  
 (B) STRAIN: Lev-12

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 678

5 TGTGCTCATT CAGGAGCTCA TTAACAACAT TGCCAAGGCT CACGGTGGTT 50  
 ACTCTGTCTT CACTGGTGTC GGTGAGCGAA CTCGTGAGGG TAACGACTTG 100  
 TACCACGAAA TGAGGGAGAC TGGTGTTATC AACCTTGAGG GTGACTCCAA 150  
 GGTGCTCTT GGTGAGTTCT TTTTCTTCTT CAGGCTAATT AGTCGATGAC 200  
 GTGGGCCCTG ACTAAAAGT TTTCTTCCAG TCTTCGGTCA GATGAACGAG 250  
 CCCCCTGGAG CCCGTGCCCCG AGTTGCCCTT ACTGGTTTGA CCATTGCCGA 300  
 GTACTTCCGT GACGAGGAAG GCCAGGATGT GTTGCTTTTC ATTGACAACA 350  
 TTTTCCGATT CACCCAGGCC GGTTCGAGG TGTCTGCCTT GCTCGGTCGT 400  
 ATCCCCCTCTG CCGTCGGTTA CCAGCCCACT CTTTCCACCG ACATGGGTGG 450  
 10 TATGCAGGAG CGAATGTAGG TTGCATTCTC TGTGATTTTA CGGCAAGCCT 500  
 TGACTTTTTT TTTCTAGTAC CACCACCAAG AAGGGTTCCA TTACCTCCGT 550  
 C 551

15

## 2) INFORMATION FOR SEQ ID NO: 679

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 552 bases  
 20 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
  
 (ii) MOLECULE TYPE: Genomic DNA  
 25 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Cryptococcus neoformans*  
 (B) STRAIN: ATCC 44104

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 679

TGTGCTCATT CAGGAGCTCA TTAACAACAT TGCCAAGGCC CACGGTGGTT 50  
 ACTCCGTCTT CACCGGTGTC GGTGAGCGAA CTCGTGAGGG TAACGACTTG 100  
 TACCACGAAA TGAGGGAGAC TGGTGTTATC AACCTTGAGG GTGACTCCAA 150  
 35 GGTGCTCTT GGTGAGTTCT TTTTCTTGGG AGCTAATTAG TCGATGACGT 200  
 GGGCCCTGAC CAAAACCGTT TCTTTCAGTC TTCGGTCAGA TGAACGAGCC 250  
 CCCTGGAGCC CGTGCCCCGAG TTGCTCTTAC TGGTTTGACC ATTGCCGAGT 300  
 ACTTCCGTGA CGAGGAAGGC CAGGATGTGT TGCTTTTCAT TGACAACATT 350  
 TTCCGATTCA CCCAGGCCGG TTCCGAGGTG TCTGCCTTGC TCGGTTCGTAT 400  
 40 CCCCTCTGCC GTCGGTTACC AGCCCACTCT TTCCACCGAT ATGGGTGGTA 450  
 TGCAAGAGCG AATGTAAGTT GCATTTTTTG TGATTTTACG GCAAGTCTTG 500  
 ACTTACATTT TTTTCTAGCA CCACCACCAA GAAGGGTTCC ATTACCTCCG 550  
 TC 552

45

## 2) INFORMATION FOR SEQ ID NO: 680

- (i) SEQUENCE CHARACTERISTICS:  
 50 (A) LENGTH: 1018 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
  
 55 (ii) MOLECULE TYPE: Genomic DNA  
  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Cunninghamella bertholletiae*  
 (B) STRAIN: ATCC 42115

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 680

```

5  TCTCCCTGCT ATTTTAAACG CTCTTGAAGT TAAGGATCAT GCTGGTGGTC 50
   GTCTTGTTCT TGAAGTTGCT CAACACTTGG GTGAAAACAC TGTACGTACT 100
   ATTGCTATGG ATGGTACTGA AGGTAAGTTT ATTTTtagAT GATCATAAAT 150
   AATTGATCAT AATGATAAAA AAAAAAGAAG AAGAAGAACA GGATGTATAT 200
   AATGGTTAAT AAATAATATT TTCATATTGK ATATAACTAT TTAATCTGTT 250
   TTTTTTCTT CATGATTATA TATATATATG TMCTAATATC TAATATGAAC 300
   CTTTTTTATA AAATTAATCA GGTCTTGTCC GTGGTCAAAA GGTTGTTGAT 350
10  ACTGGTGCTC CTATTACCAT TCCTGTTGGT AAGGAAGTTC TTGGTCGTAT 400
   CATCAACGTT ATTGGTGAAC CCATTGATGA ACGTGGTCCT ATTAAGGCCA 450
   AGTCTCACCG TGCTATTCAC GCTGAAGCTC CAGAATTCGT TGATCAATCT 500
   CCTACTCCCG AAATTCCTGA AACTGGTATT AAGGTTGTCG ATTTATTAGC 550
   TCCTTATGTA CTGGTGGTA AGATTGGTCT TTTCGGTGGT GCTGGTGTAG 600
15  GTAAAACTGT CTTGATTCAA GAACCTATTA ACAACATTGC TAAAGCCCAT 650
   GGTGGTTACT CTATTTTCTG TGGTGTGGT GAACGTACTC GTGAAGGTAA 700
   CGATTTATAC CACGAAATGA TGGAACTGG TGTCATTAAA CTTGAAGGTG 750
   ACTCCAAGTG TGCTCTTGTA TTCGGTCAAA TGAACGAACC TCCTGGTGCT 800
   CGTGCCCGTG TTGCTTTAAC TGGTTTAACC ATTGCTGAAT ACTTCCGTGA 850
20  TGAAGAAGGT CAAGATGTGT TACTTTTCAT TGATAACATT TTCCGTTTCA 900
   CTCAAGCTGG TTCTGAAGTA TCTGCCCTTT TAGGTCGTAT TCCATCTGCT 950
   GTAGGTTACC AACCCACTTT ATCTACTGAT ATGGGTGGTA TGCAAGAACG 1000
   TATTACTACT ACCAAGAA 1018

```

25

## 2) INFORMATION FOR SEQ ID NO: 681

```

30  (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 23 bases
     (B) TYPE: Nucleic acid
     (C) STRANDEDNESS: Single
     (D) TOPOLOGY: Linear

```

35 (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 681

```

40  GGISSITTYG GIISIGGIAA RAC 23

```

## 2) INFORMATION FOR SEQ ID NO: 682

```

45  (i) SEQUENCE CHARACTERISTICS:
     (A) LENGTH: 26 bases
     (B) TYPE: Nucleic acid
     (C) STRANDEDNESS: Single
     (D) TOPOLOGY: Linear

```

50

(ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 682

```

55  GTIACIGGYT CYTCRAARTT ICCICC 26

```

## 2) INFORMATION FOR SEQ ID NO: 683

60

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 683

GTIACIGGIT CISWIAWRTC ICCICC

26

2) INFORMATION FOR SEQ ID NO: 684

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 3267 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida tropicalis*  
 (C) ACCESSION NUMBER: M64984

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 684

30	ATGGCTGGTG	CTTTAGAAAA	CGCAAGAAAA	GAAATTAAAC	GTCTTTCATT	50
	AGATGACACC	AATGAATCCC	AATATGGTCA	AATCTATTCT	GTTTCCGGTC	100
	CGGTTGTTAT	TGCCGAAAAC	ATGATGGGAT	GTGCCATGTA	CGAATTGGTT	150
	AAAGTTGGTC	ATGATAATTT	AGTTGGGGAA	GTTATTAGAA	TTAATGGTGA	200
	TAAAGCAACC	ATTCAAGTTT	ATGAAGAAAC	TGCAGGGGTC	ACTGTTGGTG	250
35	ATCCAGTTTT	AAGAACTGGT	AAACCATTAT	CTGTTGAATT	AGGTCCTGGT	300
	TTAATGGAAA	CTATTTATGA	TGGTATTCAA	AGACCTTTAA	AAGCCATTAA	350
	AGATGAATCC	CAATCTATTT	ATATCCCAAG	AGGTATTGAT	GTTCTGCTT	400
	TATCAAGAAC	TGTTCAATAT	GATTTCACTC	CAGGTCAATT	GAAAGTTGGT	450
40	GATCATATCA	CTGGTGGGGA	CATTTTTGGT	TCTATTTATG	AAAACCTCTT	500
	ATTGGATGAC	CATAAGATTT	TGTTACCTCC	AAGAGCAAGA	GGTACTATTA	550
	CTTCTATTGC	TGAAGCCGGT	TCTTATAATG	TTGAAGAACC	AGTTTTGGAA	600
	GTTGAATTTG	ATGGTAAGAA	ACATAAATAC	TCTATGATGC	ATACATGGCC	650
	AGTTAGAGTT	CCAAGACCAG	TTGCTGAAAA	ATTGACTGCT	GATCATCCAT	700
45	TGTTGACCGG	TCAAAGAGTC	TTGGATTCTT	TATTCCCATG	TGTTCAAGGT	750
	GGTACTACTT	GTATCCCAGG	GGCTTTTGGT	TGTGGTAAAA	CTGTTATTTT	800
	TCAATCTTTG	TCCAAATTCT	CCAACCTCTG	TGTTATTATC	TATGTTGGTT	850
	GTTTCACTAA	AGGTACTCAA	GTCATGATGG	CTGATGGTGC	CGACAAATCT	900
	ATTGAATCTA	TTGAAGTTGG	TGACAAAGTC	ATGGGTAAAG	ATGGTATGCC	950
50	AAGAGAAGTT	GTTGGCTTAC	CAAGAGGTTA	TGATGATATG	TACAAGGTTT	1000
	GTCAACTTTT	TAGTACTAGA	CGTAATGCTA	AATCCGAAGG	CTTGATGGAT	1050
	TTCACTGTTT	CTGCTGATCA	TAAACTTATC	TTGAAAACCT	AACAAGATGT	1100
	CAAGATTGCT	ACACGTAAAA	TTGGTGGCAA	CACCTATACT	GGTGTTACTT	1150
	TCTATGTTTT	GGAAAAGACT	AAGACTGGTA	TTGAATTAGT	TAAAGCCAAG	1200
55	ACTAAAGTTT	TCGGTCATCA	TATCCATGGT	CAAAATGGCG	CTGAAGAAAA	1250
	AGCTGCTACT	TTTGCTGCTG	GCATTGACTC	TAAAGAATAC	ATTGATTGGA	1300
	TCATTGAAGC	TAGAGATTAT	GTACAAGTTG	ATGAAATTGT	CAAGACCAGC	1350
	ACCACTCAAA	TGATCAACCC	AGTTCATTTT	GAATCTGGTA	AACCTCGGTAA	1400
	CTGGTTACAC	GAACACAAGC	AAAACAAATC	ACTTGCTCCA	CAATTGGGTT	1450
60	ACTTGTTGGG	TACTTGGGCT	GGTATTGGAA	ATGTTAAATC	TTCTGCTTTC	1500

	ACCATGAACT	CCAAAGATGA	TGTTAAATTA	GCTACAAGAA	TTATGAACTA	1550
	CTCTTCAAAA	TTGGGCATGA	CTTGTTCTTC	TACTGAATCC	GGTGAATCA	1600
	ATGTCGCTGA	AAACGAAGAA	GAATTTTTC	ATAACCTTG	TGCTGAAAAG	1650
	GATGAAGCTG	GTGATTTCAC	TTTTGATGAA	TTTACCGATG	CTATGGATGA	1700
5	ATTGACTATC	AATGTTTCATG	GTGCAGCTGC	AAGCAAGAAG	AACAATTTGT	1750
	TGTGGAATGC	TTTGAAATCT	CTTGGTTTCA	GAGCCAAGTC	TACTGATATT	1800
	GTCAAGAGTA	TTCTCAACA	TATTGCTGTT	GATGATATTG	TTGTCAGAGA	1850
	ATCTTTGATT	GCCGGTTTAG	TTGATGCTGC	TGGTAATGTT	GAAACCAAAT	1900
	CCAATGGTTC	TATTGAAGCT	GTTGTTAGAA	CTTCTTTCAG	ACATGTCGCT	1950
10	AGAGGTCTTG	TCAAGATTGC	TCATTCTTTG	GGTATTGAAT	CATCTATTAA	2000
	TATTAAAGAT	ACTCACATTG	ATGCTGCTGG	TGTTAGACAA	GAATTTGCTT	2050
	GTATTGTCAA	TTTGACTGGT	GCTCCACTTG	CTGGTGTTCT	TTCTAAATGT	2100
	GCACTTGCAA	GAAACCAAAC	TCCAGTTGTC	AAATTTACCA	GAGACCCAGT	2150
	TTTGTTCAAC	TTTGATTTGA	TCAAATCTGC	AAAAGAAAAC	TATTATGGTA	2200
15	TTCTTTTGGC	TGAAGAAACT	GATCATCAAT	TCCTTTTATC	CAACATGGCC	2250
	TTGGTGCACA	ACTGTGGTGA	ACGTGGTAAT	GAGATGGCTG	AAGTTTTGAT	2300
	GGAATTCCCA	GAATTGTTTA	CTGAAATTTT	TGGTAGAAAA	GAACCAATTA	2350
	TGAAACGTAC	CACTTTGGTT	GCCAATACTT	CTAATATGCC	AGTCGCTGCC	2400
	AGAGAAGCTT	CTATTTATAC	TGGTATTACA	TTGGCTGAAT	ATTTCAGAGA	2450
20	TCAAGGTAAG	AATGTTTCTA	TGATTGCTGA	TTCTTCTTCA	CGTTGGGCTG	2500
	AAGCTTTGAG	AGAAATTTCT	GGTAGATTGG	GTGAAATGCC	TGCTGATCAA	2550
	CGTTTCCCAG	CTTATTTGGG	TGCTAAATTT	GCTTCTTTCT	ATGAGCGTGC	2600
	CGGTAAAGCC	ACTGCTTTGG	GTTCAACAGA	TAGAGTTGGT	TCAGTTTCTA	2650
	TTGTTGCTGC	TGTTTCTCCA	GCTGGTGGTG	ATTTCTCTGA	TCCAGTTACT	2700
25	ACTTCTACTT	TGGGTATTAC	TCAAGTTTTC	TGGGGGTTGG	ATAAGAAATT	2750
	GGCCCAAAGA	AAACATTTCC	CATCTATTAA	CACCAAGTGT	TCTTATTCTA	2800
	AATACACCAA	TGTTTTGAAC	AAATACTATG	ATTCCAACCT	TCCAGAATTC	2850
	CCACAATTGA	GAGACAAAAT	TAGAGAAATT	TTATCTAATG	CTGAAGAATT	2900
	GGAACAAGTT	GTCAATTAG	TTGGTAAATC	TGCATTGTCT	GATTCTGATA	2950
30	AGATTACTTT	AGATGTTGCT	ACCTTGATTA	AAGAAGATTT	CTTGCAACAA	3000
	AATGGTTATT	CTTCATATGA	TGCATTCTGT	CCAATTTGGA	AGACTTTTGA	3050
	TATGATGAGA	GCATTTATTT	CATATTATGA	TGAAGCACAA	AAAGCAATTG	3100
	CCAATGGTGC	TCAATGGTCT	AAATTAGCTG	AAAGTACTAG	TGATGTTAAA	3150
	CATGCTGTTT	CTTCAGCTAA	ATTCTTTGAA	CCATCAAGAG	GTCAAAAAGA	3200
35	AGGTGAAAAA	GAATTTGGAG	ATTTATTAAC	CACTATCTCC	GAAAGATTTG	3250
	CTGAAGCTTC	AGAATAA				3267

## 40 2) INFORMATION FOR SEQ ID NO: 685

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1782 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 50 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus hirae*  
 (B) STRAIN: ATCC 9790  
 (C) ACCESSION NUMBER: D17462

## 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 685

	TTGCAAATTG	GAAAAATCAT	AAAAGTCTCC	GGTCCTCTCG	TTATGGCAGA	50
	AAATATGTCA	GAAGCAAGTA	TTCAAGACAT	GTGTTTAGTG	GGAGATTTAG	100
	GAGTCATCGG	CGAAATCATT	GAGATGCGTC	AAGATGTGGC	GTCTATTCAA	150
60	GTATATGAAG	AAACTTCAGG	AATTGGTCCC	GGAGAACCTG	TTCGTTCCAC	200

	TGGGGAAGCA	CTATCTGTTG	AGCTAGGACC	AGGAATCATT	TCACAAATGT	250
	TTGACGGGAT	TCAAAGACCA	CTGGATACAT	TTATGGAAGT	GACTCAAAGT	300
	AACTTCTTAG	GACGTGGGGT	CCAATTACCA	GCTTTAGATC	ATGAGAAACA	350
	ATGGTGGTTT	GAAGCGACAA	TGAAGAAGG	AACAGAAGTA	AGTGCTGGAG	400
5	ACATCATTTG	GTACGTGGAT	GAAACGAAGA	TCATTAGCA	CAAAATCATG	450
	GTCCCTAATG	GTATCAAAGG	AACTGTACAA	AAAATTGAAT	CTGGATCATT	500
	TACGATCGAT	GATCCGATTT	GTGTGATCGA	AACGGAACAA	GGCTTAAAAG	550
	AGCTGACGAT	GATGCAAAAA	TGGCCAGTAC	GTCTGGTTCG	ACCAATCAAA	600
	CAAAAAATTAA	ATCCAGATGT	ACCGATGATC	ACCGGTCAAA	GGGTCATTGA	650
10	CACGTTTTTC	CCAGTAACTA	AAGGAGGAGC	GGCAGCCGTT	CCAGGTCCGT	700
	TTGGTGCAGG	GAAGACAGTT	GTGCAACACC	AGATTGCTAA	GTGGTCGGAC	750
	GTAGATCTAG	TGGTTTACGT	TGGTTGTGGG	GAACGAGGAA	ATGAAATGAC	800
	GGATGTCGTC	AATGAATTTT	CTGAACTGAT	CGATCCAAAT	ACAGGCGAGT	850
	CTTTGATGGA	ACGAACGTG	TTGATCGCTA	ATACATCGAA	CATGCCAGTA	900
15	GCTGCTCGAG	AAGCTTCTAT	TTATACGGGA	ATCACGATTG	CCGAGTACTT	950
	CCGTGACATG	GGGTATGATG	TAGCAATCAT	GGCAGATTCC	ACTTCTCGTT	1000
	GGGCAGAAGC	ACTGCGTGAA	ATGAGCGGAC	GTTTAGAAGA	AATGCCTGGT	1050
	GATGAAGGTT	ATCCCCTTA	TCTGGGCTCT	CGTCTAGCTG	AATACTATGA	1100
	ACGTTCAGGA	CGTGTCATTG	CTCTAGGCTC	TGACCAACGT	GAGGCGAGTA	1150
20	TCACTGCCAT	CAGTGCAGTT	TCTCCTTCTG	GTGGAGATAT	CTCTGAACCA	1200
	GTGACTCAA	ATACCTTACG	TGTGGTGAAG	GTTTTCTGGG	GATTAGATTC	1250
	TAGTCTTGCT	CAAAAAAGAC	ATTTTCCATC	GATTAAGTGG	ATCCAAAGTT	1300
	ACTCATTATA	TTCAACAGAA	GTTGGCAGAT	ATATGGATCA	AATCTTACAA	1350
	CAGGATTGGT	CTGATATGGT	AACTGAAGGC	ATGCGGATCT	TGCAAGAAGA	1400
25	AGAACAATTA	AATGAAAATCG	TGCGCTTGGT	AGGGATCGAT	TCGCTTTCTG	1450
	ATAACGATCG	CTTGACCCTT	GAAGTAGCAA	AATCGATTTCG	AGAAGACTAT	1500
	TTACAACAAA	ATGCTTTTGA	TGATGTAGAT	ACGTTTACTT	CAAGAGAAAA	1550
	ACAATTCAAC	ATGTTGAAAG	TTATTTTGAC	TTTGGGAAA	GAAGCTCGAA	1600
	AAGCCTTATC	GTTGGGAGCG	TACTTCAATG	AAATCATGGA	AGGTACAGTA	1650
30	GCGGTCAGAG	AACGCATTAG	TCGGAGCAAG	TATATTCCAG	AAGAAGAGTT	1700
	AGCCAAAATT	AGTAGTATAA	ATGAAGAAAT	CAAAGAAACG	ATCCAATTGA	1750
	TTGTTTCAGA	AGGAGGGATG	ACCGATGATT	AA		1782

35

## 2) INFORMATION FOR SEQ ID NO: 686

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1781 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

45

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Chlamydia pneumoniae*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 686

50	CAGGGACATG	TTATAGAAGC	TTATGGAAAC	TTGTTACGTG	TACGCTTTGA	50
	CGGATATGTT	AGACAAGGTG	AAGTTGCATA	TGTCAACGTA	GATAATACCT	100
	GGTTAAAAGC	AGAAGTGATT	GAAGTTGCTG	ATCAAGAAGT	CAAGGTTTCAG	150
	GTATTTGAAG	ATACACAAGG	CGCGTGTCTG	GGAGCTCTTG	TTACGTTTTTC	200
55	AGGACATCTT	TTAGAAGCCG	AGTTAGGGCC	TGGCTTGCTT	CAGGGCATTT	250
	TCGATGGACT	TCAAATTCGT	CTTGAGGTGC	TAGCTGAAGA	TAGTTCTTTC	300
	TTGCAGAGAG	GCAAGCATGT	TAATGCTATT	TCTGATCATA	ATTTATGGAA	350
	TTATACTCCC	GTAGCTTCTG	TTGGGGATAC	TTTAAGACGA	GGAGATCTTC	400
	TAGGAACAGT	ACCTGAAGGA	CGATTTACTC	ATAAGATTAT	GGTTCCTTTT	450
60	TCTTGCTTTC	AAGAGGTTAC	CCTGACTTGG	GTAATTTCTG	AAGGAACCTA	500

	TAATGCTCAT	ACTGTGGTTCG	CAAAAGCTCG	AGATGCTCAG	GGTAAAGAAT	550
	GTGCC'TTTAC	TATGGTGCAA	AGATGGCCGA	TCAAACAAGC	TTTTATTGAA	600
	GGAGAGAAGA	TCCCTGCGCA	TAAGATTATG	GATGTGGGTT	TGCGAATCTT	650
	AGATACGCAA	ATTCCAGTAT	TGAAGGGGGG	AACTTTCTGT	ACCCCAGGAC	700
5	CTTTTGGTGC	AGGGAAAACA	GTCTTACAAC	ACCATCTTTC	TAAGTACGCT	750
	GCTGTAGATA	TTGTGATTTT	GTGTGCGTGC	GGAGAGCGTG	CTGGTGAAGT	800
	TGTTGAGGTA	TTACAAGAGT	TCCCTCATCT	TATCGACCCC	CATACCGGAA	850
	AGTCTTTAAT	GCACAGAACA	TGTATTATTT	GTAACACATC	ATCCATGCCT	900
	GTGGCTGCCC	GAGAGTCTTC	GATCTATTTA	GGAGTGACGA	TTGCAGAATA	950
10	CTATCGCCAG	ATGGGACTAG	ATATTCTGCT	TTTAGCTGAT	TCTACATCCC	1000
	GATGGGCACA	AGCCCTTAGA	GAGATTTCCG	GACGTCTTGA	AGAAATCCCT	1050
	GGAGAGGAAG	CATTTCCCTGC	ATACCTGTCT	TCTAGAATAG	CTGCTTTTTA	1100
	TGAGCGAGGA	GGAGCTATCA	CCACGAAAGA	TGGTTCTGAA	GGATCTTTAA	1150
	CTATATGTGG	TGCGGTGTCT	CCTGCAGGAG	GAAACTTTGA	AGAACCAGTC	1200
15	ACTCAATCTA	CATTAGTGTG	ATCGGAGCG	TTCTGTGGTC	TTTCAAAGC	1250
	ACGAGCTGAC	GCACGTAGGT	ATCCTTCAAT	AGACCCTTTG	ATTTCTTGGT	1300
	CAAAATATTT	GAACCAGGTA	GGACAAATTT	TAGAAGAGAA	GGTTTCAGGC	1350
	TGGGGTGGTG	CTGTGAAAAA	AGCAGCACAG	TTTCTAGAGA	AAGGTTTCAGA	1400
	AATCGGCAAG	CGTATGGAAG	TTGTCACTGA	AGAAGGGGTT	TCTATGGAAG	1450
20	ACATGGAAAT	CTACTTAAAG	GCAGAACTTT	ATGATTTTGT	TkATCTCCAG	1500
	CAGAACGCAT	TCGATCCTGT	GGACTGTTAT	TGTCCTTTTG	AGAGACAGAT	1550
	AGAGTTATTT	TCATTAATCA	GTCGTATTTT	TGATGCTAAA	TTTGT'TTTTG	1600
	ATAGTCCTGA	TGATGCAAGA	AGCTTTT'TCC	TTGAGCTGCA	GAGCAAGATT	1650
	AAGACATTAA	ATGGCCTGAA	ATTTCTTTCA	GAGGAATATC	ATGAGAGTAA	1700
25	AGAGGTCATA	GTTAGACTGT	TGGAAAAAAC	AATGGTACAA	ATGGCGTAAG	1750
	GATATGCAAA	CAATCTACAC	AAAATAACT	G		1781

## 30 2) INFORMATION FOR SEQ ID NO: 687

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1758 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Halobacterium salinarum*  
 (C) ACCESSION NUMBER: S56356

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 687

45	ATGAGTCAAG	CTGAAGCAAT	CACTGACACC	GGCGAAATCG	AGAGCGTGAG	50
	CGGTCCGGTC	GTGACCGCCA	CGGGCCTCGA	CGCGCAGATG	AACGACGTGC	100
	TCTACGTGGG	CGACGAGGGT	CTGATGGGCG	AGGTCATCGA	GATCGAAGGC	150
	GACGTAACCA	CCATCCAGGT	CTACGAGGAG	ACGTCCGGCA	TCGGGCCGGG	200
50	CCAGCCCGTC	GACAACACGG	GCGAACCCTG	CACCGTGGAC	CTGGGGCCCGG	250
	GGATGCTGGA	CTCCATCTAC	GATGGTGTGC	AGCGTCCACT	GGACGTCCCTC	300
	GAAGACGAGA	TGGGGGCGTT	CCTCGATCGC	GGTGT'TGACG	CACCCGGCAT	350
	CGACCTCGAC	ACCGACTGGG	AGTTCGAGCC	CACCGTCGAG	GCGGGCGACG	400
	AGGTCGCGGC	CGGCGATGTC	GTCGGCACCG	TCGACGAAAC	GGTCAGCATC	450
55	GAACACAAGG	TTCTGGTGCC	CCCCCGCAGT	GACGGCGGCG	AAGTCGTCGC	500
	CGTCGAATCC	GGCACGTTCA	CCGTCGACGA	CACGGTCGTC	GAGTTGGACA	550
	CCGGCGAGGA	GATCCAGATG	CACCAAGAGT	GGCCGGTCCG	CCGCCAGCGC	600
	CCCACCTCG	ACAAGCAGAC	GCCGACGGAG	CCGCTGGTGT	CCGGCCAGCG	650
	CATCCTCGAC	GGCCTGTTCC	CGATCGCGAA	AGGCGGGACG	GCCGCGATCC	700
60	CGGGGCCGTT	CGGGTCCGGG	AAGACGGTCA	CCCAGCAGTC	CCTCGCGAAG	750



	TTCGCCGACG	CGGACATCGT	TGTCTACATC	GGCTGTGGTG	AGCGCGGCAA	800
	CGAGATGACG	GAAGTCATCG	AGGACTTCCC	GGAGCTGCCC	GACCCCCAGA	850
	CCGGGAACCC	GCTGATGGCC	CGCACCACGC	TCATCGCGAA	CACGTCTGAAC	900
	ATGCCGGTTG	CCGGGCGTGA	GTCTCTGCATT	TACACGGGAA	TCACCATCGC	950
5	GGAGTACTAC	CGCGACATGG	GCTACGACGT	GGCGCTGATG	GCCGACTCCA	1000
	CCTCGCGGTG	GGCGGAGGCC	ATGCGGGAGA	TCTCCTCGCG	ACTCGAGGAG	1050
	ATGCCCGGCG	AGGAGGGGTA	TCCCGCGTAC	CTGGCCGCCC	GCCTCTCGGA	1100
	GTTCTACGAG	CGCGCCGGCT	ACTTCGAGAA	CTTCAACGGG	ACCGAGGGCT	1150
	CCATCTCGGT	CATCGGTGCG	GTGTCGCCCG	CCGGCGGGGA	CTTCTCCGAG	1200
10	CCGGTCACCC	AGAACACGCT	GCGCATCGTG	AAGACGTTCT	GGGCGCTTGA	1250
	CTCGGACCTC	GCCGAGCGCC	GGCACTTCCC	GGCGATCAAC	TGGGACGAGT	1300
	CCTACAGCCT	CTACAAGGAC	CAACTCGACC	CGTGGTTCAC	GGACAACGTC	1350
	GTCGACGACT	GGGCCGAGCA	GCGCCAGTCG	GCGGTCGACA	TCCTCGACGA	1400
	GGAATCCGAA	CTCGAAGAGA	TCGTGCAGCT	CGTCGGGAAG	GACGCGCTGC	1450
15	CCGAGGACCA	GCAGCTCACG	CTGGAAGTCG	CGCGGTACAT	CCGCGAGGCG	1500
	TGGCTCCAGC	AGAACGCGCT	CCACGACGTG	GATCGCTACT	GCCCGCCCGA	1550
	GAAGACGTAC	GCCATCCTCT	CCGGCATCAA	GACGCTTCAC	GAGGAGTCCT	1600
	TTGAGGCGTT	GGACGCCGGT	GTGCCAGTCG	AGGAGATCAC	GTCGATCGAC	1650
	GCCGCGCCGC	GCCTGAATCG	TCTCGGCACG	ACGCCCAGCG	ACGAGCACGA	1700
20	GGCGGAGGTC	GCGGAGATCA	AACAGCAGAT	TACCGAGCAG	CTTCGGGAGC	1750
	TCTACTGA					1758

25 2) INFORMATION FOR SEQ ID NO: 688

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 3118 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

35 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Human  
 (C) ACCESSION NUMBER: L09234

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 688

40	GAATTCGGGC	AGCTGACTAG	TCTTGTGATT	GGGGTCCTGG	GCTGATAAAA	50
	TCATTCCAAA	TGACGAGCAC	ATTGATAAAG	ACGTCCGATG	AGGACCGGGA	100
	GTCCAAATTC	GGCTTTGTTT	TTGCCGTATC	TGGACCTGTG	GTGACAGCTG	150
	AACGAATGGC	CGGTTCTGCT	ATGTACGAAC	TGGTGCGTGT	CGGTTATTAT	200
45	GAAGTGGTCG	GAGAGATCAT	CCGGTTGGAG	GGTGACATGG	CAACAATCCA	250
	AGTATACGAA	GACACCTCAG	GTGTGACAGT	AGGCGATCCC	GTGCTGCGCA	300
	CAGGCAAGCC	GCTGTCCGTG	GAAGTGGGAC	CCGGAATCAT	GGGCAGCATC	350
	TTCGACGGTA	TCCAGCGACC	GCTGAAGGAT	ATCAATGAAC	TGTCAAATAG	400
	TATCTACATC	CCGAAAGGTG	TCAATGTGCC	TGCCCTGAGT	CGCACTGCAC	450
50	AGTGGGACTT	CAGTCCCCTG	AGTGTCAAGG	TTGGAAGCCA	CATTACTGGT	500
	GGTGACCTGT	ACGGTTTGTT	CCACGAAAAT	ACTCTGGTGA	AACACAAGTT	550
	GCTGCTGCCG	CCCCGTGCCA	AGGGAAGTGT	CACGTACATT	GCAGAACCTG	600
	GAAACTACAC	AGTTGATGAT	GTTGTCCTGG	AGACAGAATT	TGACGGCGAG	650
	CGATCAAAGT	TCACCATGCT	GCAAGTGTGG	CCTGTACGTC	AGCCCAGGCC	700
55	TGTTACAGAA	AAGTTGCCAG	CTAACTACCC	CCTCCTTACT	GGCCAGCGTG	750
	TGCTCGACTC	CCTATTCCCG	TGTGTCCAGG	GTGGAACAAC	AGCTATTCTT	800
	GGGGCCTTCG	GATGTGGCAA	GACTGTAATA	TCACAGTCTT	TGTCAAATAA	850
	CTCAAATCC	GATGTAATTA	TCTATGTAGG	TTGTGGTGAG	CGAGGTAATG	900
	AAATGTCAGA	AGTACTCAGG	GATTTCCCGC	AGTTGTCGTT	GGAGATTGAT	950
60	GGTGTGACTG	AATCAATCAT	GAAGAGAACA	GCCCTGGTCG	CAAACACATC	1000

	AAACATGCCT	GTGGCTGCTC	GAGAAGCATC	TATCTACACA	GGTATTACAC	1050
	TGTCAGAAATA	CTTCAGGGAC	ATGGGTTACA	ATGTATCCAT	GATGGCTGAC	1100
	TCAACTTCAC	GATGGGCCGA	AGCTCTTCGA	GAAATCTCAG	GTGATTGGC	1150
	TGAAATGCCT	GCCGACAGCG	GTTATCCCGC	CTACCTAGGT	GCACGACTTG	1200
5	CCAGTTTCTA	CGAGCGTGCC	GGCCGTGTGA	AGTGCTTGGG	TAACCCAGAC	1250
	AGGGAGGGCT	CCGTGAGTAT	AGTGGGCGCC	GTGTCGCCGC	CCGGTGGAGA	1300
	CTTCTCAGAT	CCCGTGACGA	CGGCCACACT	AGGTATCGTC	CAGGTGTTCT	1350
	GGGGTCTCGA	CAAGAAACTT	GCCCAGCGAA	AGCACTTCCC	ATCCATCAAC	1400
	TGGCTCATCT	CGTAGAGTAA	ATACATGCGT	GCTCTGGATG	ACTTCTACGA	1450
10	CAAGAATTTT	CCAGAGTTTG	TCCCCTGCG	TACAAAGGTG	AAGGAGATTT	1500
	TGCAGGAGGA	AGAAGACCTG	TCTGAAATTG	TGCAGTTGGT	CGGTAAAGCT	1550
	TCATTGGCAG	AAACTGACAA	GATCACACTT	GAGGTTGCCA	AACTATTAAA	1600
	GGATGATTTT	CTGCAACAGA	ACAGCTATTC	ACCATATGAC	CGTTTCTGCC	1650
	CATTCTACAA	GACAGTAGGA	ATGCTGAAAA	ATATGATTGC	TTTCTACGAT	1700
15	ATGTCTCGGC	ATGCAGTTGA	ATCTACTGCT	CAGAGCGAGA	ACAAGATCAC	1750
	TTGGAATGTT	ATTAGAGATT	CTATGGGCAA	TATTCTGTAT	CAGCTTTCCT	1800
	CCATGAAATT	CAAGGATCCA	GTCAAGGATG	GAGAAGCGAA	GATCAAGGCA	1850
	GACTTTGAGC	AGCTTCATGA	AGACATTGAG	CAAGCCTTCA	GGAACCTGGA	1900
	GGATTAAAGT	GGTAGCTGCC	AGTGGTTCTC	TCGGTGCAGT	TGTCACATTT	1950
20	GGCAAGCTCT	GTAGGGTTGC	CGAGTGGCAT	CGGTGCTAGA	CACCTGAGCA	2000
	TTCCTTTGCC	ACATAAAGAC	TAAAGCAGGT	GGAATTTTCA	TTGTAAAAAG	2050
	CTGGTTCCAT	TGGTGCCTAAG	ATTATGTTGT	GCCCTTTTCT	GCTTCTCACA	2100
	TTCCAACAGA	GGAAATTTACT	TCCAGTTTTC	TTCCATTTTC	CTCCTCATTT	2150
	TAAGTGTCGG	TACAGAGGCA	ATAATCTGAT	AACTCTGTAC	CGTCACTTAC	2200
25	AAGCAGGGAG	AATTTGTAAT	TATTACAAAT	CCCATTATCT	CTGTGCACCA	2250
	CAGCCTTGTA	AATTCATTTG	TCCCAGGACT	CCCTCTTGTC	TGTACGTGAG	2300
	ATTGCCGTCT	GTATGTATGT	ACACACCGTA	CTGCAGTATT	TGAAGTCAGT	2350
	CAGAAGGTGA	ATTACACCAC	TTACTCATTG	TGTCACGTAG	CAAGTGTGCA	2400
	AACTGCCATC	CATTGTCCTA	TTTATTCACA	TAAGTAGTTT	TCTTTGCATT	2450
30	TCCAGTGTTG	CAAATTGTGT	TTAGAAAATT	ATGCCATCGA	GACTGGTCTGA	2500
	ACCTCACATT	GTAATCTCAGT	ATTTACACAC	ACGTTTACTT	GCTACAGAAA	2550
	TGTAGAAAAA	ATAATTGTTG	TATATTGAAA	GTACAAGTGA	CAAAGTTGCA	2600
	TTTAAAATGG	TGAATGTATT	TTATATTTCT	TTTGTAGACA	CAAGAGTTAA	2650
	TGCATTTTGC	TTAATGGAGA	TGTATGTAAA	CCTAAAATAG	CAGTTTGTGC	2700
35	ACAAATTATG	TATATGTGAA	ATGGAGATGG	TTTCTAATTT	GCTGATTGAT	2750
	TGCCAGTATT	AATTTAAACA	ACTGTAGTTG	TGGGATGTAG	TGGGAAGATT	2800
	TTTTTTTTCC	TATAAAATTG	GTGGATGTAT	GTGTCGGAGA	TTTTGATTGT	2850
	ATGTGTAAAA	TAGTGATCCC	AGTAACTGTA	AAGCTTTAGA	ATACAGTTAC	2900
	TGACTGTATA	GTTGTACAGG	TGTTGTTACT	TTTAAGAATT	TATTGACACA	2950
40	AAGGTGAAAG	TCTATTATTG	TATTGTAAATG	TTTAAAGCAT	TTAAGGTTTA	3000
	AAAATCCTAC	TTCTGTGTAT	AAATGTTACC	ATTCCTCATA	TAACATAACT	3050
	GTGTAGAAAT	ACAGTCAACT	TCATGTTTCA	TAGCATTTCA	CTGTTGTCAC	3100
	ATAAATTATG	CCCGGAAT				3118

45

## 2) INFORMATION FOR SEQ ID NO: 689

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 1836 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 55 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 60 (A) ORGANISM: *Plasmodium falciparum*  
 (B) STRAIN: 3D7  
 (C) ACCESSION NUMBER: L08200

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 689

	ATGACAAAAG	TTGCTGTTGA	AAAAGAGGAA	CCAGGAGTTG	TTTATAAGGT	50
5	GGCTGGTTCA	TTAGTTATTG	CTGAGAATAT	GAGTGGAAC	CGTATGTACG	100
	AGTTAGCTAA	AGTAGGATGG	AATAAATTGG	TTGGAGAAAT	TATTAGATTA	150
	GAAGGGAATT	ATGCATATAT	ACAAGTTTAT	GAAGATACTT	CAGGTTTATC	200
	TGTAGGAGAC	CCTGTTATAA	AAACAGGAAA	TGCTTTATCA	GTCGAATTGG	250
	GTCCTGGTAT	TTTAGATAAT	ATTTATGATG	GTATTCAAAG	ACCATTAGAA	300
10	AGAATAGCAA	ATGTGTGTGG	TGATGTATAT	ATATATAAAG	GTATTGATAT	350
	GACATCTTTA	GATCATGATA	AACAATGGCA	ATTTTATGCT	GATAAGAAAT	400
	TAAAATTAAA	TGATATTGTT	ACTGGTGGAG	ATATCTTTGG	ATTTGTTGAT	450
	GAAAATAAAT	TATTTAAAGA	ACACAAAATT	ATGGCTCCAC	CTAATGCTAA	500
	AGGGAGGCTT	ACATATATTG	CTCCAGATGG	ATCATATACT	TTAAAAGATA	550
15	AAATATTTGA	ATTAGAATAT	CAAGGAAAAA	AATATACATA	TGGTTTATCT	600
	CATTTATGGC	CTGTTCGTGA	TCCTAGACCT	GTTTTAGAAA	AGGTAACAGG	650
	GGATACTTTA	TTATTAACAG	GGCAAAGAGT	TTTAGATTCC	TTATTTCCAA	700
	CAGTTCAAGG	AGGTACTTGT	GCTATTCCCT	GTGCATTGG	TTGTGGAAAA	750
	ACTTGTGTTT	CTCAGGCCTT	ATCAAAATAT	TCTAATAGTG	AAGTTATTAT	800
20	ATATGTAGGA	TGTGGTGAAA	GAGGTAATGA	AATGGCTGAA	ATTTTATCCG	850
	ACTTTCCTGA	ATTAACACT	AAAGTAGATA	ATGAAGATGT	AGGTATTATG	900
	CAAAGAACGT	GTTTAGTTGC	TAATACTTCT	AACATGCCTG	TCGCTGCAAG	950
	AGAAGCTAGT	ATTTATACAG	GTATTACTTT	ATGTGAATAT	TTCCGTGATA	1000
	TGGGTTATAA	TGCTACCATG	ATGGCTGATA	GTACAAGTAG	ATGGGCAGAA	1050
25	GCCTTAAGAG	AAATTTTCAGG	ACGTTTAGCT	GAAATGCCTG	CAGATAGTGG	1100
	TTATCCAGCT	TATTTAGGTG	CTAGATTAGC	TTCTTTTAT	GAACGTGCAG	1150
	GAAAAGTCAA	ATGTATTGGT	TCTCCATCTC	GTATAGGATC	CATTACAATT	1200
	GTGGGTGCTG	TGTCTCCACC	AGGTGGTGAT	TTCTCTGACC	CTGTAACTAC	1250
	AGCAACCATG	TCTATTGTTC	AAGCATTTTG	GGGGTTAGAT	AAAAAACTAG	1300
30	CTCAAAGAAA	ACATTTCCCT	TCTGTTAATT	GGTCTACATC	CTTTTCAAAG	1350
	TATGTCAGAC	AATTAGAACA	ATACTTTGAT	AATTTTGATC	AAGATTTCTT	1400
	ATCTTTAAGA	CAAAAAATTA	GTGATATTTT	ACAACAAGAA	AGTGACTTGA	1450
	ATGATATTGT	TCAACTAGTA	GGAAAGGATT	CATTATCAGA	AGACCAAAAA	1500
	GTTGTTATAG	AAGTAGCCAA	AATTATTAGA	GAAGATTTTC	TTCAACAAAA	1550
35	TGCATTTAGC	GATTATGATT	ATATGTGCCC	ATTACAAAAA	ACAGTTGGTA	1600
	TGATGAGAAT	TATTTGCCAC	TTTTATGCTC	AATGCTTAAG	AACATTACAA	1650
	GAATATGACT	CAAGAGAAAG	AAAAATTGGT	TGGGGATCTA	TATATAATAC	1700
	ATTAAGACCA	ACTATAAATA	AAATTACACA	TATGAAATTT	GAAAACCCAA	1750
	AAAATTCAGA	TGAATATTTT	AAAAGTATT	TTAAGGCACT	TGAAGAAGAA	1800
40	ATAACAGTAG	GTTTAAGAAA	CTTGATGGAA	AAATGA		1836

## 2) INFORMATION FOR SEQ ID NO: 690

45

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3216 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

50

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

55

- (A) ORGANISM: *Saccharomyces cerevisiae*
- (B) STRAIN: X2180-1A
- (C) ACCESSION NUMBER: J05409

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 690

60

	ATGGCTGGTG	CAATTGAAAA	CGCTCGTAAG	GAAATAAAAA	GAATCTCATT	50
	AGAAGACCAT	GCTGAATCTG	AATATGGTGC	CATCTATTCT	GTCTCTGGTC	100
	CGGTCGTCAT	TGCTGAAAT	ATGATTGGTT	GTGCCATGTA	CGAATTGGTC	150
	AAGGTCGGTC	ACGATAACCT	GGTGGGTGAA	GTCATTAGAA	TTGACGGTGA	200
5	CAAGGCCACC	ATCCAAGTTT	ACGAAGAAAC	TGCAGGCCTT	ACGGTCGGTG	250
	ACCCTGTTTT	GAGAACAGGT	AAGCCTCTGT	CGGTAGAATT	GGGTCCTGGT	300
	CTGATGGAAA	CCATTTACGA	TGGTATTCAA	AGACCTTTGA	AAGCCATTAA	350
	GGAAGAATCG	CAATCGATTT	ATATCCCAAG	AGGTATTGAC	ACTCCAGCTT	400
	TGGATAGGAC	TATCAAGTGG	CAATTTACTC	CGGGAAAGTT	TCAAGTCGGC	450
10	GATCATATTT	CCGGTGGTGA	TATTTACGGT	TCCGTTTTTG	AGAATTCGCT	500
	AATTTCAAGC	CATAAGATTC	TTTTGCCACC	AAGATCAAGA	GGTACAATCA	550
	CTTGGAATTG	TCCAGCTGGT	GAGTACACTT	TGGATGAGAA	GATTTTGGAA	600
	GTTGAATTTG	ATGGCAAGAA	GTCTGATTTT	ACTCTTTACC	ATACTTGGCC	650
	TGTTCTGTGT	CCAAGACCAG	TTACTGAAAA	GTTATCTGCT	GACTATCCTT	700
15	TGTTAACAGG	TCAAAGAGTT	TTGGATGCTT	TGTTTCCTTG	TGTTCAAGGT	750
	GGTACGACAT	GTATTCCAGG	TGCTTTTGGT	TGTGGTAAGA	CCGTTATCTC	800
	TCAATCTTTG	TCCAAGTACT	CCAATTCTGA	CGCCATTATC	TATGTCGGGT	850
	GCTTTGCCAA	GGGTACCAAT	GTTTTAATGG	CGGATGGGTC	TATTGAATGT	900
	ATTGAAAACA	TTGAGGTTGG	TAATAAGGTC	ATGGGTAAAG	ATGGCAGACC	950
20	TCGTGAGGTA	ATTAATTTGC	CCAGAGGAAG	AGAACTATG	TACAGCGTCG	1000
	TGCAGAAAAG	TCAGCACAGA	CCCCACAAAA	GTGACTCAAG	TCGTGAAGTG	1050
	CCAGAATTAC	TCAAGTTTAC	GTGTAATGCG	ACCCATGAGT	TGGTTGTTAG	1100
	AACACCTCGT	AGTGTCCGCC	GTTTGTCTCG	TACCATTAAG	GGTGTGCAAT	1150
	ATTTTGAAGT	TATTACTTTT	GAGATGGGCC	AAAAGAAAGC	CCCCGACGGT	1200
25	AGAATTGTTG	AGCTTGTCAG	GGAAGTTTCA	AAGAGCTACC	CAATATCTGA	1250
	GGGGCCTGAG	AGAGCCAACG	AATTAGTAGA	ATCCTATAGA	AAGGCTTCAA	1300
	ATAAAGCTTA	TTTTGAGTGG	ACTATTGAGG	CCAGAGATCT	TTCTCTGTTG	1350
	GGTTCCCATG	TTCGTAAAGC	TACCTACCAG	ACTTACGCTC	CAATTCTTTA	1400
	TGAGAATGAC	CACTTTTTCG	ACTACATGCA	AAAAAGTAAG	TTTCATCTCA	1450
30	CCATTGAAGG	TCCAAAAGTA	CTTGCTTATT	TACTTGGTTT	ATGGATTGGT	1500
	GATGGATTGT	CTGACAGGGC	AACTTTTTTCG	GTTGATTCCA	GAGATACTTC	1550
	TTTGATGGAA	CGTGTTACTG	AATATGCTGA	AAAGTTGAAT	TTGTGCGCCG	1600
	AGTATAAGGA	CAGAAAAGAA	CCACAAGTTG	CCAAAACGTG	TAATTTGTAC	1650
	TCTAAAGTTG	TCAGAGGTAA	TGGTATTCGC	AATAATCTTA	ATACTGAGAA	1700
35	TCCATTATGG	GACGCTATTG	TTGGCTTAGG	ATTCTTGAAG	GACGGTGTCA	1750
	AAAATATTCC	TCTTTTCTTG	TCTACGGACA	ATATCGGTAC	TCGTGAAACA	1800
	TTTCTTGCTG	GTCTAATTGA	TTCTGATGGC	TATGTTACTG	ATGAGCATGG	1850
	TATTAAAGCA	ACAATAAAGA	CAATTCATAC	TTCTGTCAGA	GATGGTTTGG	1900
	TTTCCCTTGC	TCGTTCTTTA	GGCTTAGTAG	TCTCGGTTAA	CGCAGAACCT	1950
40	GCTAAGGTTG	ACATGAATGG	CACCAAACAT	AAAATTAGTT	ATGCTATTTA	2000
	TATGTCCTGGT	GGAGATGTTT	TGCTTAACGT	TCTTTCGAAG	TGTGCCGGCT	2050
	CTAAAAAATT	CAGGCCTGCT	CCCGCCGCTG	CTTTTGCACG	TGAGTGCCGC	2100
	GGATTTTATT	TCGAGTTACA	AGAATTGAAG	GAAGACGATT	ATTATGGGAT	2150
	TACTTTATCT	GATGATTCTG	ATCATCAGTT	TTTGCCTTGCC	AACCAGGTTG	2200
45	TCGTCCATAA	TTGCGGAGAA	AGAGGTAATG	AAATGGCAGA	AGTCTTGATG	2250
	GAATTTCCAG	AGTTATATAC	TGAAATGAGC	GGTACTAAAG	AACCAATTAT	2300
	GAAGCGTACT	ACTTTGGTCG	CTAATACATC	TAACATGCCG	GTTGCAGCCA	2350
	GAGAAGCTTC	TATTTACACT	GGTATCACTC	TTGCAGAATA	CTTCAGAGAT	2400
	CAAGGTAAAA	ATGTTTCTAT	GATTGCAGAC	TCTTCTTCAA	GATGGGCTGA	2450
50	AGCTTTGAGA	GAAATTTCTG	GTCGTTTGGG	TGAGATGCCT	GCTGATCAAG	2500
	GTTTCCCAGC	TTATTTGGGT	GCTAAGTTGG	CCTCCTTTTA	CGAAAGAGCC	2550
	GGTAAAGCTG	TTGCTTTAGG	TTCCCCAGAT	CGTACTGGTT	CCGTTTCCAT	2600
	CGTTGCTGCC	GTTTCGCCAG	CCGGTGGTGA	TTTCTCAGAT	CCTGTTACTA	2650
	CTGCTACATT	GGGTATCACT	CAAGTCTTTT	GGGGTTTAGA	CAAGAAATTG	2700
55	GCTCAAAGAA	AGCATTTCCC	ATCTATCAAC	ACATCTGTTT	CTTACTCCAA	2750
	ATACACTAAT	GTCTTGAACA	AGTTTATATG	TTCCAATTAC	CCTGAATTTC	2800
	CTGTTTTAAG	AAGTCGTATG	AAGGAAATTC	TATCAAACGC	TGAAGAATTA	2850
	GAACAAGTTG	TTCAATTAGT	TGGTAAATCG	GCCTTGCTCTG	ATAGTGATAA	2900
	GATTACTTTG	GATGTTGCCA	CTTTAATCAA	GGAAGATTTT	TTGCAACAAA	2950
60	ATGGTTACTC	CAC TTATGAT	GCTTTCTGTC	CAATTTGGAA	GACATTTGAT	3000

	ATGATGAGAG	CCTTCATCTC	GTATCATGAC	GAAGCTCAAA	AAGCTGTTGC	3050
	TAATGGTGCC	AACTGGTCAA	AACTAGCTGA	CTCTACTGGT	GACGTTAAGC	3100
	ATGCCGTTTC	TTCATCTAAA	TTTTTTGAAC	CAAGCAGGGG	TGAAAAGGAA	3150
	GTCCATGGCG	AATTCGAAAA	ATTGTTGAGC	ACTATGCAAG	AAAGATTGTC	3200
5	TGAATCTACC	GATTAA				3216

## 2) INFORMATION FOR SEQ ID NO: 691

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1860 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Schizosaccharomyces pombe*
  - (B) STRAIN: 972 h-
  - (C) ACCESSION NUMBER: S47814

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 691

25	ATGGCGGGAG	GAATTGAACT	GGCCAAGAAG	GCTATCAGGA	GCCTCAAAAA	50
	TTACGACGAG	CATGAAAACC	GATATGGATC	TATTTTCAGC	GTTTCTGGTC	100
	CTGTCGTTGT	TGCAGCCAAT	ATGCTTGGAT	GTTTCGATGTA	CGAACTCGTT	150
	CGCGTTGGTC	ATGAAGAACT	AGTTGGTGAA	GTAATTCGTA	TCCATCAAGA	200
30	TAAATGTACT	ATTCAAGTAT	ACGAAGAGAC	GTCCGGTCTC	ACTGTTGGTG	250
	ATCCTGTCCA	ACGCACTGGA	AAGCCATTAT	CTGTTGAATT	AGGTCCTGGT	300
	TTAGCTGAGA	CTATTTATGA	TGGTATCCAA	CGTCCGTTAA	AGCAAATTTT	350
	CGACAAATCT	CAAAGTATTT	ATATTCCTAG	AGGTATTAAT	ACAGAATCAC	400
	TTAATCGTGA	GCATAAGTGG	GATTTACACAC	CAAATAAGGA	TTTACGCATT	450
35	GGCGATCATG	TATCCGGTGG	TGATGTTTTT	GGTTCGTAT	TTGAAAACCTC	500
	TCTTTTCAAT	GATCATAAAA	TTATGTTACC	CCCTAGAGCC	CGTGGTACCG	550
	TCACATATAT	TGCTGAAGCT	GGATCATACC	ATGTTGATGA	AAAACCTTCTT	600
	GAAGTCGAGT	TTAATGGCAA	GAAACATTCT	TTTAGTATGT	TGCATACTTG	650
	GCCTGTCCGT	GCTGCTCGTC	CAGTTGCGGA	CAACTTAACT	GCTAATCAAC	700
40	CTTTATTGAC	TGGTCAACGT	GTTTTGGATG	CGTTATACCC	CTGTGTTCAA	750
	GGTGCGACTA	CTGCTATCCC	CGGTGCCTTT	GGTTGTGGTA	AAACAGTTAT	800
	TTCACAAATCT	CTTTCTAAGT	ACTCTAATTC	TGATTTGATT	GTTTACGTCG	850
	GTTGTGGTGA	ACGTGGAAC	GAAATGGCAG	AAGTGTTAAT	GGATTTCCCA	900
	GAACATAACAA	TTGATATTAA	TGGTAAACCA	GAGCCCATTA	TGAAGCGTAC	950
45	TACATTGGTA	GCCAACACTT	CTAACATGCC	TGTCGCTGCT	CGTGAAGCTT	1000
	CCATTTATAC	CGGTATTACA	CTTGCTGAAT	ATTATCGTGA	TCAAGGTAAG	1050
	AACGTTTCAA	TGATGGCTGA	TTCTACATCT	CGTTGGGCTG	AAGCTTTGCG	1100
	TGAAATTTCT	GGTCGTTTGG	CTGAGATGCC	TGCCGATTCT	GGTTATCCCG	1150
	CTTATTTGGG	TGCCAAATG	GCTTCTTTTT	ACGAACGTGC	TGGTCGTGCT	1200
50	CGTTGCTTGG	GAAGTCCTGA	CCGTGAAGGA	ACAGTTTCAA	TTGTTGGAGC	1250
	TGTTTCTCCT	CCGGGTGGTG	ATTTTCTGA	TCCTGTTACT	AGTGCAACCT	1300
	TGGGAATTGT	TCAAGTCTTC	TGGGGTTTGG	ACAAGAAATT	GGCCCAACGT	1350
	AAACACTTTC	CCTCAATCAA	CACCTCTCTT	TCCTATTCTA	AATACATCAA	1400
	TGCTTTGCAA	CCTTGGTATG	AGGAAAGAGT	TCCAGGCTTT	AATACTCTTC	1450
55	GTGATCAAAT	CAAACAGATC	ATTCAACAAG	AAGATTCCAT	GTTGGAAATT	1500
	ATTCAAGTTGG	TTGGTAAGTC	GGCTCTTTCT	GAAACGGATA	AAGTTACTTT	1550
	GGACATAGCC	GGTATTATTA	AGAATGACTT	CTTACAACAA	AACGGTTATT	1600
	CTGATTACGA	TCGCTGTTGC	CCTCTTTACA	AGACTTATCA	TATGATGCGA	1650
	AACATGATTG	CTTACTACAC	AAAGGCTAAA	AGTGCCGTTG	AAACTGGTAG	1700
60	CGTTCCTTGG	TCAAAGATTA	AAGAAAGTAC	TTCAGATATC	TTTTATGAGT	1750

TAACCTCGAT	GAAATTCGAA	AACCCTAATG	AAGGCGAGAA	GGAAATAGTC	1800
GAACACTATG	AAACTCTGCA	CAAGAAGATT	GAGGACAAGT	TTCACACTCT	1850
GA CTGAGTAA					1860

5

## 2) INFORMATION FOR SEQ ID NO: 692

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1833 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma congolense*  
 (B) STRAIN: IL3000  
 20 (C) ACCESSION NUMBER: Z25814

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 692

25	ATGACGAGCG	ATAAAAACCC	TTACAAAACA	GAGCAGCGCA	TGGGGGCCGT	50
	GAAGGCCGTC	TCCGGGCCAG	TTGTCAATTGC	TGAAAACATG	GGCGGTAGCG	100
	CTATGTATGA	GCTTGTGCAG	G TAGGTTCTT	TCCGGTTAGT	GGGCGAGATC	150
	ATTCGTCTAG	AGGGCGATAC	CGCCACTATT	CAGGTCTATG	AGGAAACAGG	200
	TGGCCTCACT	GTCCGAGACC	CGGTGTACTG	TACGGGTAAG	CCTCTTTCGC	250
	TTGAGCTTGG	ACCTGGAATC	ATGCTTGAAA	TATTTGACGG	TATCCAGCGG	300
30	CCTCTTGACA	CCATCTACCG	CATGGTGGAA	AACGTGTTTA	TCCCCAGGGG	350
	CGTTCAGGTG	AAGTCACTCA	ATGACCAGAA	ACAGTGGGAC	TTTAAGCCAT	400
	GCCTGAAGGT	TGGAGATCTT	GTGTCTGGTG	GTGATATCAT	TGGCTCAGTG	450
	GTGGAGAACT	CTCTCATGTA	CAATCACAGC	ATTATGATTC	CGCCCAATGT	500
	GCGGGGCCGT	GTTACTTCCA	TTGTTCTTTC	AGGAAATTAC	ACCCTCCAAG	550
35	ATGACATTAT	TGAATTGGAA	TATAATGGGA	CAGTGAAATC	ACTAAAACCT	600
	ATGCATCGCT	GGCCAGTACG	GACCCGCGCT	CCTGTGGCGT	CAAAAGAATC	650
	CGGCAATCAT	CCGCTTCTCA	CCGGACAGCG	TGTGCTCGAT	GCTCTCTTTC	700
	CATCCGTCCA	GGGTGGAACA	TGCGCCATCC	CTGGCGCGTT	TGGATGCGGA	750
	AAGACGGTTA	TCAGTCAGGC	TCTTTCGAAG	TTCTCCAACA	GCGACGCTGT	800
40	TATCTATGTC	GGCTGCGGCG	AGCGTGGGAA	TGAGATGGCA	GAGGTGCTCA	850
	TGGACTTCCC	CACACTCACC	ACCGTTATTG	ATGGTCGTGA	GGAGTCCATC	900
	ATGAAGCGTA	CCTGCCTGGT	GGCAAACACC	TCAAATATGC	CTGTCGCTGC	950
	TCGTGAGGCG	TCTATTTACA	CTGGCATCAC	TTTAGCTGAG	TATTATCGTG	1000
	ATATGGGCAA	GCACATTGCT	ATGATGGCCG	ACTCTACCTC	TCGATGGGCT	1050
45	GAGGCTCTCC	GTGAGATCTC	TGGGCGTCTC	GCTGAAATGC	CCGCTGATGG	1100
	TGGTTACCTT	GCGTACCTCA	GTGCGCGTCT	TGCTTCCTTC	TACGAGCGTG	1150
	CGGGGCGCGT	GACATGCATC	GGTGGGCCAA	AACGCGAGGG	CTCAGTAACC	1200
	ATCGTTGGTG	CCGTTTCTCC	TCCTGGAGGT	GACTTTTCTG	ACCCAGTGAC	1250
	GTCCGCTACG	CTTGGTATTG	TGCAAGTCTT	TTGGGGTCTT	GAGAAGCGTC	1300
50	TTGCGCAACG	TAAACACTTT	CCTTCTGTTA	ATTGGCTCAT	TTCTTATTCA	1350
	AAATACCTTA	ATGCTTTGGA	GCCCTTCTTC	AACACGCTTG	ACCCTGACTA	1400
	CATGCGCCTG	CGGTCAGTTG	CTGCGGAGAT	CCTTCAGCGT	GAGGAAGAGT	1450
	TGCAAGAAAT	TGTTCAACTT	GTCCGGTAAG	ACTCACTTTC	GGAGTCTGAC	1500
	AAAATTTATC	TAGAAACGGC	TAAGGTTATT	CGTGAAGAGT	TTCTCCAGCA	1550
55	GAATGCCTTT	ACGCCGTACG	ACAAGTATTG	CCCGCCGTAC	AAGACCTGCT	1600
	GGATGCTACG	TAACATTGTC	GCGTTCTACG	AGGAGAGCCA	GCGCGTTGTA	1650
	GCTGAGTCCG	CTGGGGAAC	TAAGATTACG	TGGAAC	TTCGTGAAAT	1700
	GATTCCTCAT	ATTTACACGG	GTTTAACTGA	GATGAAGTTC	CGTGATCCTC	1750
	AGGAGGGTGA	GGAGGCCAAC	GTAGAATTCT	ACAGAAAACA	AAATGAGGAA	1800
60	ATTGTCAGCG	CATTCGCCTC	GCTGCTGCAA	TAA		1833

## 2) INFORMATION FOR SEQ ID NO: 693

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1758 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Thermus thermophilus*  
 (B) STRAIN: HB8  
 (C) ACCESSION NUMBER: D63799

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 693

20	ATGATCCAAG	GGGTGATCCA	GAAGATCGCG	GGCCCCGGCGG	TGATCGCCAA	50
	GGGCGATGCTC	GGGGCCCCGCA	TGTACGACAT	CTGCAAGGTG	GGCGAAGAGG	100
	GCCTCGTG	CGAGATCATC	CGCCTGGACG	GGGACACGGC	CTTCGTCCAG	150
	GTCTACGAGG	ACACCTCGGG	CCTAAAGGTG	GGGGAGCCCG	TGGTCTCCAC	200
25	GGGCCTTCCC	TTGGCGGTGG	AGCTCGGCCC	CGGGATGCTG	AACGGCATCT	250
	ACGACGGCAT	CCAGCGCCCC	CTGGAGCGCA	TCCGGGAGAA	GACGGGGATC	300
	TACATCACCC	GGGGCGTGGT	GGTCCACGCC	CTGGACCGGG	AGAAGAAGTG	350
	GGCCTGGACG	CCCATGGTCA	AGCCCCGGGA	CGAGGTGCGG	GGGGGTATGG	400
	TCCTGGGCAC	GGTGCCCGAG	TTCGGCTTCA	CCCACAAGAT	CCTGGTACCC	450
30	CCGGACGTGC	GGGGCCGGGT	CAAGGAGGTG	AAGCCCGCCG	GGGAGTACAC	500
	CGTGGAGGAG	CCGGTGGTGG	TCCTCGAGGA	CGGCACCGAG	CTCAAGATGT	550
	ACCACACCTG	GCCCGTTTCG	CGGGCGAGGC	CCGTGCAAAG	GAAGCTTGAC	600
	CCCAACACCC	CCTTCCTCAC	GGGGATGCGC	ATCCTGGACG	TCCTCTTCCC	650
	CGTGGCCATG	GGGGGCACCG	CCGCCATCCC	TGGGGCCCTT	GGCAGCGGCA	700
35	AGACCGTGAC	CCAGCAGTCC	CTGGCCAAGT	GGTCCAACGC	CGACGTGGTG	750
	GTCTACGTGG	GCTGCGGGGA	GCGGGGGAAC	GAGATGACCG	ACGTGCTCGT	800
	GGAGTTCCCC	GAGCTCACCG	ACCCCAAGAC	GGGTGGGCCC	TTGATGCACC	850
	GCACCGTCCT	CATCGCCAAC	ACCTCCAACA	TGCCCCGTGG	CGCCCCGCGAG	900
	GCCAGCATCT	ACGTGGGCGT	GACCATCGCC	GAGTACTTCC	GCGACCAGGG	950
40	CTTCTCCGTG	GCCCTCATGG	CCGACTCCAC	GAGCCGCTGG	GCCGAGGCTT	1000
	TGCGCGAGAT	CTCTAGCCGC	CTCGAGGAGA	TGCCCCGCCA	GGAGGGCTAC	1050
	CCGCCCTACC	TCGCCGCCAG	GCTCGCCGCC	TTCTACGAGC	GGGCGGGCAA	1100
	GGTCATCACC	CTGGGCGGCG	AGGAGGGGGC	GGTGACCATC	GTGGGGGCGG	1150
	TCTCCCCGCC	GGGCGGCGAC	ATGTCCGAGC	CCGTGACCCA	GTCCACCTTG	1200
45	AGGATCGTGG	GGGCCTTCTG	GCGGCTTGAC	GCCTCCCTGG	CCTTCCGCCG	1250
	CCACTTCCCC	GCCATCAACT	GGAACGGCTC	CTACAGCCTC	TTCACCTCCG	1300
	CCCTTGACCC	CTGGTACCGG	GAGAACGTGG	CCGAGGACTA	CCCCGAGCTC	1350
	CGCGACGCCA	TCTCCGAGCT	TTTGACGCGG	GAGGCGGGCC	TCCAGGAGAT	1400
	CGTCCAGCTC	GTGGGGCCCG	ACGCCCTCCA	GGACGCCGAG	CGCCTCGTCA	1450
50	TTGAGGTGGG	CCGGATCATC	CGCGAGGACT	TCCTGCAGCA	GAACGCCTAC	1500
	CACGAGGTGG	ACGCCTACTG	CTCCATGAAG	AAGGCCTACG	GGATCATGAA	1550
	GATGATCCTC	GCCTTCTACA	AGGAGGCGGA	GGCGGCCATC	AAGCGGGGGG	1600
	TTTCCATAGA	CGAGATCCTG	CAGCTCCCCG	TTCTGGAGCG	CATCGGCCCG	1650
	GCCCCGTACG	TGAGCGAGGA	GGAGTTCCCC	GCCTACTTTG	AGGAGGCCAT	1700
55	GAAGGAGATC	CAGGGGGCCT	TCAAGGCTGG	CCTAAAGGGG	GAGAGATGGA	1750
	CCTTCTGA					1758

## 60 2) INFORMATION FOR SEQ ID NO: 694

600

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 694

CGGCGCIATC YTSGTTGTTG C

21

2) INFORMATION FOR SEQ ID NO: 695

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 695

GTTTCACGTG ATGACGTACA

20

2) INFORMATION FOR SEQ ID NO: 696

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 696

ATIGGICAYR TIGAYCAYGG IAARAC

26

2) INFORMATION FOR SEQ ID NO: 697

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 697

CCIACIGTIC KICCRCCYTC RCG

23



## 2) INFORMATION FOR SEQ ID NO: 698

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1185 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (C) ACCESSION NUMBER: extracted from J01690

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 698

20	GTGTCTAAAG	AAAAATTTGA	ACGTACAAAA	CCGCACGTTA	ACGTTGGTAC	50
	TATCGGCCAC	GTTGACCACG	GTAAACTAC	TCTGACCGCT	GCAATCACCA	100
	CCGTACTGGC	TAAAACCTAC	GGCGGTGCTG	CTCGTGCATT	CGACCAGATC	150
	GATAACGCGC	CGGAAGAAAA	AGCTCGTGGT	ATCACCATCA	ACACTTCTCA	200
	CGTTGAATAC	GACACCCCGA	CCCGTCACTA	CGCACACGTA	GA CTGCCCCG	250
25	GGCACGCCGA	CTATGTTAAA	AACATGATCA	CCGGTGCTGC	TCAGATGGAC	300
	GGCGCGATCC	TGGTAGTTGC	TGCGACTGAC	GGCCCGATGC	CGCAGACTCG	350
	TGAGCACATC	CTGCTGGGTC	GTCAGGTAGG	CGTTCCGTAC	ATCATCGTGT	400
	TCCTGAACAA	ATGCGACATG	GTTGATGACG	AAGAGCTGCT	GGA ACTGGTT	450
	GAAATGGAAG	TTCGTGAACT	TCTGTCTCAG	TACGACTTCC	CGGGCGACGA	500
30	CACTCCGATC	GTTTCGTGGT	CTGCTCTGAA	AGCGCTGGAA	GGCGACGCAG	550
	AGTGGAAGC	GAAAATCCTG	GA ACTGGCTG	GCTTCCTGGA	TTCTTATATT	600
	CCGGAACCAG	AGCGTGCGAT	TGACAAGCCG	TTCTTGCTGC	CGATCGAAGA	650
	CGTATTCTCC	ATCTCCGGTC	GTGGTACCGT	TGTTACCGGT	CGTGTAGAAC	700
	GCGGTATCAT	CAAAGTTGGT	GAAGAAGTTG	AAATCGTTGG	TATCAAAGAG	750
35	ACTCAGAAGT	CTACCTGTAC	TGGCGTTGAA	ATGTTCCGCA	AACTGCTGGA	800
	CGAAGGCCGT	GCTGGTGAGA	ACGTAGGTGT	TCTGCTGCGT	GGTATCAAAC	850
	GTGAAGAAAT	CGAACGTGGT	CAGGTACTGG	CTAAGCCGGG	CACCATCAAG	900
	CCGCACACCA	AGTTCGAATC	TGAAGTGATC	ATTCTGTCCA	AAGATGAAGG	950
	CGGCCGTCAT	ACTCCGTTCT	TCAAAGGCTA	CCGTCCGCAG	TTCTACTTCC	1000
40	GTACTACTGA	CGTGA CTGGT	ACCATCGAAC	TGCCGGAAGG	CGTAGAGATG	1050
	GTAATGCCGG	GCGACAACAT	CAAAATGGTT	GTTACCCTGA	TCCACCCGAT	1100
	CGCGATGGAC	GACGGTCTGC	GTTTCGCAAT	CCGTGAAGGC	GGCCGTACCG	1150
	TTGGCGCGGG	CGTTGTTGCT	AAAGTTCTGG	GCTAA		1185

## 2) INFORMATION FOR SEQ ID NO: 699

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 699

GTIACIGGYT CYTYRARRTT ICCICC

## 2) INFORMATION FOR SEQ ID NO: 700

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
10 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 700

15 TIRTIGAYGT CGARTTCCCT CARG

24

## 2) INFORMATION FOR SEQ ID NO: 701

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 701

30 GTGTTACGA TCATCGATGC G \_\_\_\_\_

21

## 35 2) INFORMATION FOR SEQ ID NO: 702

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 702

CTCTCGATAT CCGCGAAGCG

20

## 50 2) INFORMATION FOR SEQ ID NO: 703

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
55 (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 703

TATGGAAATT CGAAACATCT

20

5

2) INFORMATION FOR SEQ ID NO: 704

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 704

AGTGCTCCAA TTAATGTTGG

20

20

2) INFORMATION FOR SEQ ID NO: 705

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

30 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 705

35 GTACAGTTCC AATACCTGAA

20

2) INFORMATION FOR SEQ ID NO: 706

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 706

50 TGAAATCTTC ACATCCAACA

20

55 2) INFORMATION FOR SEQ ID NO: 707

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single

60

604

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 707

TWACCATTTC AGTACCTTCT GTAA

25

10

2) INFORMATION FOR SEQ ID NO: 708

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

15

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 708

TCRTCCATIC CIARIATIGC IATAT

26

25

2) INFORMATION FOR SEQ ID NO: 709

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 1656 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

35

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Borrelia burgdorferi*

(C) ACCESSION NUMBER: extracted from AE001122

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 709

	ATGAATGAAG	TTTTATTTGT	AAAGACTGCT	GGTAGGAATT	TAAAAGCAGA	50
	AGTAATTCGT	ATTAGGGGCA	ATGAAGTTGA	TGCACAGGTT	TTTGAATTGA	100
45	CAAAAGGGAT	ATCTGTTGGA	GACCTAGTTG	AATTTACAGA	CAAACCTTTTA	150
	ACAGTTGAAC	TCGGACCAGG	GCTTTTAACT	CAAGTATATG	ATGGGCTTCA	200
	AAATCCTTTG	CCTGAATTGG	CTATTCAATG	TGGATTTTTT	TTAGAAAGGG	250
	GAGTATATTT	AAGGCCCTTG	AATAAAGATA	AAAAGTGGAA	TTTTAAAAAA	300
	ACCTCCAAAG	TTGGAGATAT	CGTTATTGCA	GGAGATTTTT	TAGGTTTTGT	350
50	AATTGAGGGA	ACTGTTCAAC	ATCAAATAAT	GATTCCATTT	TATAAAAGGG	400
	ATTCTTATAA	AATTGTGGAG	ATTGTAAGTG	ATGGCGACTA	TTCGATTGAT	450
	GAGCAAATTG	CTGTAATTGA	AGATGATTCT	GGTATGAGGC	ATAATATTAC	500
	AATGTCTTTT	CATTGGCCTG	TTAAAGTTCC	TATTACTAAT	TATAAGGAAC	550
	GCCTTATTCC	TAGTGAACCT	ATGTTGACTC	AAACTAGAAAT	TATAGATACA	600
55	TTTTTCCCAG	TTGCCAAAGG	TGGAACTTTT	TGCATTCCGG	GTCCTTTTGG	650
	AGCAGGAAAA	ACGGTCTCTC	AGCAGGTTAC	AAGTCGAAAT	GCTGATGTTG	700
	ATGTAGTGAT	TATTGCAGCT	TGTGGTGAGC	GAGCAGGAGA	AGTGGTAGAA	750
	ACTCTTAAAG	AATTTCCCGA	ATTAATGGAT	CCAAAAACCG	GCAAACTCTT	800
	AATGGACAGG	ACTTGTATTA	TTTGTAATAC	ATCTTCAATG	CCAGTTGCAG	850
60	CTAGAGAAGC	TTCTGTTTAT	ACTGCTATTA	CTATTGGTGA	GTATTACAGG	900

```

CAAATGGGCC TTGATATTCT TCTTTTGGCA GATTCAACTT CAAGATGGGC      950
TCAAGCAATG AGAGAAATGT CTGGACGCCT TGAGGAAATT CCTGGCGAGG      1000
AGGCTTTTCC GGCATATCTT GAGTCTGTTA TTGCTTCCTT TTATGAAAGG      1050
GCAGGTATTG TAGTCTTTAA TAATGGGGAT ATTGGATCTG TAACAGTTGG      1100
5  TGGCTCTGTA AGTCCTGCTG GTGGTAATTT TGAAGAGCCA GTTACTCAAG      1150
CAACTTTAAA AGTTGTAGGA GCATTTACAG GGCTTACAAG AGAAAGGTCT      1200
GATGCTAGGA AATTTCACAG TATTAGTCCCT CTTGAATCTT GGAGTAAATA      1250
TAAAGGCGTT ATTGATCAAA AAAAGACTGA ATATGCAAGA TCTTTTTTGG      1300
TGAAAGGTAA TGAAATTAAT CAAATGATGA AAGTTGTGG AGAAGAAGGC      1350
10 ATAAGTAACG ATGATTTTTT AATTTATTTA AAATCCGAGC TACTTGATTC      1400
GTGCTATTGG CAGCAAAATT CATTGATTC TATTGATGCT GCTGTTAGTT      1450
CAGAGCGTCA AAATTATATG TTTGATATAG TTTATAACAT TCTTAAACT      1500
AACTTTGAGT TTTCTGATAA ACTTCAAGCA AGAGATTTTA TAAATGAGTT      1550
AAGGCAAAAT CTTTTAGACA TGAATCTTTC TTCTTTTAAG GATCATAAGT      1600
15 TTAATAAATT GGAGCATGCT TTGGGTGAAT TGATAAATTT TAAAAAGGTA      1650
ATTTAG

```

## 20 2) INFORMATION FOR SEQ ID NO: 710

## (i) SEQUENCE CHARACTERISTICS:

```

25 (A) LENGTH: 1818 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

```

----- (A) ORGANISM: Treponema pallidum
        (B) STRAIN: Nichols
        (C) ACCESSION NUMBER: extracted from AE000520

```

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 710

```

GTGATCAAAG ACGATGTGGT TACAGGCCGT GTAGTGAGGG TGTCTGGTCC      50
CATTGTGTAT GCCGAGGGCC TCTCTGCGTG CAGCGTATAC GATGTTGTCTG      100
ACGTAGGGGA AGCATCGCTC ATCGGAGAAA TTATCCGGTT GGATGAGAGC      150
40 AAGGCGGTG TGCAAGTATA CGAGGATGAC ACAGGTATGC GAGTCGGGGA      200
GAAGGTGACA AGCTTGCGTC GACCACTCTC AGTCCGCTTA GGGCCTGGAT      250
TAATCGGCAC CATTATGAC GGTATTCAGC GCCCACTTGA GCGCCTCTTC      300
CAAGAAGACG GCGCCTTCTT GCGTCCTGGT GCGCGTTTAC AACCGCTTGA      350
TGGCTCCGTA CGCTGGGATT TTCGTCCTCA TTGTAACGAG CGCGGTGAGG      400
45 CCCTGTGCGC GGGGATTCCG ATTGCACCTG GGTCACTGTT AGGGACCGTG      450
CAGGAGACTC CTTCTGTTGT GCACACTATC ATGGTTCCTC CTGACATCCG      500
GGGGAGCGTG CTATCTTCGT TCAAGGGCGC AGGTGCTTAC ACAATAGATG      550
AAGAAATTGG ACGCACTGAT CTTGGTGAGC CGCTTTTTCT ATCCCACTAC      600
TGGCCAGTGC GTCGTGCGCG TCCTTTCAGC AAAAACTTG CAGTGTGTGA      650
50 GCCACTAGTT ACTGGACAGC GGGCGATTGA TGTTTTCTTC CCCCTATCAA      700
AGGGAGGAAC GGCGGCTATT CCAGGGGGAT TTGGAAGTGG GAAGACAATG      750
ACGCAGCATG CCGTTGCCAA GTGGTGTGAT GCAGATATTA TCGTGACAT      800
CGGCTGCGGA GAGCGGGGCA ACGAGATGAC AGACGTGCTC TCTGAATTTT      850
CCAAACTCAT CGATCCGCGC ACAGGACGCT CTCTTATGGA GCGGACGATT      900
55 TTGATCGCAA ATACGTCCAA TATGCCTGTG TCCGCACGCG AGGTGTCGCT      950
GTATTCAGGG ATTACCCTTG CGGAATACTA CCGTGATATG GGTATGCATG      1000
TGGCCATCAT GGCTGATTCT ACCAGCCGCT GGGCGGAGGC GCTGCGTGAA      1050
TTGTCTGGGC GCATGGAAGA AATGCCTGCG GAGGAGGGAT TCCCTGCGTA      1100
CCTTCCGACG CGTCTTGCA G AATTTTATGA GCGCGCAGGA CGCGTGAAA      1150
60 CCTGTGTGGC GCGCGAGGGC TCTGTGAGCA TCATTGGTGC TGTTTCTCCC      1200

```

	CTGGGTGGAG	ATTTCTCTGA	GCCGGTGACG	CAGCACACAA	AGCGCTTCAT	1250
	CCGTTGCTTT	TGGGCCTTGG	ATCGTGAAC	TGCACACGCG	CGTCATTACC	1300
	CTGCCATTGG	GTGGATAGAT	TCATACTCTG	AATATGCGCA	GGAAGTAAGT	1350
	GCATGGTGGA	GTAAGTATGA	CCC GCGCGCA	GGCGCGTTGC	GCGCCGCAGC	1400
5	CTTGGATTGT	CTGAGAAAGG	AACAGCGGTT	ACAGCAAATT	GTCAGGCTTG	1450
	TCGGTCCTGA	TGCGCTGCCT	GGAGAAGATC	GTCTGGTGCT	AATGGTGTGT	1500
	GAAATGATCA	AAGGTGGCTT	TCTGCAGCAG	AACGCTTTTG	ATCCGACGGA	1550
	TGTGTTCTCC	TGTCCCGAAA	AGCAGGTGCA	GATCTTGCGT	ACCATAGTGG	1600
	ATTTTACGA	ACGTGCCGTG	GTGCTGCTGC	GTGCAGGTAT	TTCGCTTTCT	1650
10	GCGCTGTCCC	AGCTTTCGTG	CCGGGAGCTC	ATCGTACGTA	TGAAAACACTAC	1700
	GTACGGGAAT	GAGGATGTAC	ACAAGATGCA	GAAAGTGTAC	GACACGATGT	1750
	GCACTGAGTT	TGACCAACTG	AGTGTGTGTG	CTGCCGCGCG	CACACAAGGG	1800
	GGGGAGAAAG	TCGAATGA				1818

15

## 2) INFORMATION FOR SEQ ID NO: 711

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1779 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Chlamydia trachomatis*

(B) STRAIN: MoPn

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 711

	CAGGGCTATG	TCGTAGAAGC	TTACGGAAAT	TTATTGCGGG	TGCATTTTGA	50
	TGGGCATGTG	CGTCAAGGAG	AAGTGGCCTA	TGTCAGCGTG	GATGATACTT	100
35	GGTTGAAAGC	GGAAATTATA	GAAGTTGTGG	GAGATGAGGT	TAAAGTCCAA	150
	GTTTTTGAGG	AAACTCAAGG	AATTTCTCGA	GGCGCTTTGG	TAACCTTTTC	200
	CGGGCATTTA	TTAGAAGCGG	AACTTGGGCC	CGGTCTATTG	CAAGGTATTT	250
	TTGACGGACT	TCAGAATCGC	TTAGAGGTAT	TGGCAGATAC	AAGCTTGTTT	300
	TTGAAAAGAG	GGGAGTATGT	TAATGCCATT	TGTCGGGAAA	CTGTATGGGC	350
40	TTATACGCAA	AAGGCTTCTG	TCGGGGATGT	TCTATCTCGG	GGAGATGTGC	400
	TTGGTACAGT	AAAGGAAGGG	CGGTTTGATC	ATAAAATCAT	GGTTCCTTTC	450
	TCTTGTTTTG	AGGAAGTGAC	TATCACTTGG	GTCATTTCTT	CAGGAGATTA	500
	CACTGTTGAT	ACCGTTATTG	CTAAAGGACG	TACTGCTTCA	GGAGCCGAGC	550
	TTGAATTTAC	AATGGTTCAG	AAATGGCCCA	TTAAACAGGC	TTTTTTTAGAA	600
45	GGGGAAAAGG	TACCGTCTCA	TGAAATTATG	GATGTTGGGT	TACGAGTATT	650
	AGATACTCAG	ATCCCCGTCT	TAAAGGGAGG	AACTTTTTGT	ACTCCAGGGC	700
	CTTTTGGTGC	AGGAAAGACC	GTTTTACAGC	ACCATTTATC	TAAGTATGCA	750
	GCTGTAGATA	TCGTAGTTTT	GTGTGCTTGT	GGAGAGCGAG	CTGGAGAGGT	800
	TGTAGAAATT	CTTCAGGAGT	TCCCGCATTT	GACAGATCCT	CATACGGGGC	850
50	AGTCTTTGAT	GCATAGGACC	TGTATTATTT	GTAATACATC	TTCCATGCCT	900
	GTAGCAGCTA	GAGAGTCCTC	CATTTATTGT	GGTATTACTA	TAGCAGAATA	950
	TTACCGTCAA	ATGGGGTTGC	ATGTTTTGTT	ATTGGCTGAC	TCGACATCTA	1000
	GATGGGCTCA	AGCTTTAAGG	GAAATTTTCA	GGCGATTAGA	AGAAATCCCT	1050
	GGAGAAGAAG	CTTTCCCAGC	CTATTTGGCG	TCTCGAATAG	CAGCTTTTTA	1100
55	TGAGCGAGGC	GGGGCTGTGA	AAATGAAAGA	TGGATCGGAA	GGCTCCTTGA	1150
	CTATCTGTGG	AGCGGTTTCT	CCCGCAGGAG	GAAATTTTGA	AGAGCCTGTT	1200
	ACACAAGCAA	CTTTATCTGT	TGTTGGGGCT	TTCTGTGGGC	TTTCTAAGGC	1250
	TAGAGCAGAT	GCTAGACGGT	ATCCTTCTAT	TGATCCGATG	ATTTTCATGGT	1300
	CTAAGTACTT	GGATTCTGTG	GCGGAGATTT	TGAGAAAAAA	AGTTCCAGGA	1350
60	TGGGGAGATT	CCGTTAATAA	AGCTTCTCGT	TTCTTAGAAG	AAGGAGCAGA	1400

AATTGGTAAG CGAATAGAAG TTGTTGGGGA AGAAGGGATT TCTATGGAAG 1450  
 ATATAGAAAT CTTTTTGAAG TCAGAGTTGT ATGATTCTG TTACTTACAG 1500  
 CAAAACGCTT TCGATGCAGA GGAAGTTTAT TGTCCCTTTG ATCGTCAAAT 1550  
 AGAGCTTTTT TCTTTAATGA GTCATATTTT TAGCTCTAGA TTCTGTTTTG 1600  
 5 ATTGTCCAGA TAATGCTCGG AGTTTCTTTT TAGAGCTTCA AAGTAAAATT 1650  
 AAAACGCTGA ATGGTCAAAA ATTCCTTTCT GAAGACTATC AGAAGGGGCT 1700  
 AGAAGTGATC TATAAACTAT TAGAAAGCAA AATGGTGCAG ACGGCGTAGG 1750  
 TATGCAAACA ATATATACAA GAATTACGG 1779

10

## 2) INFORMATION FOR SEQ ID NO: 712

## (i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 965 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecalis*  
 (B) STRAIN: V583

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 712

GTGCAAAATTG GAAAAATTGT CAAAGTTTCA GGTCCCTTTGA TTTTAGCTGA 50  
 AAACATGTCA GATGCTAGTA TCCAAGACAT TTGTCATGTA GGAGATTTAG 100  
 30 GCGTTATCGG AGAGATTATT GAAATGCGAG GCGACGTCGC TTCGATTCAA 150  
 GTATATGAAG AAACAACAGG CATTGGACCA GGAGAACCAG TTATTTCAAC 200  
 AGGAGAACCA TTATCTGTTG AATTAGCCCC AGGTTTAATT GCCGAAATGT 250  
 TTGATGGTAT TCAACGACCA TTGGATACAT TTCAAGAAGT AACCCACAGT 300  
 AACTTTTTTAG GCCGTGGCGT TAAAATTGAT GCGTTAGATC GTGAGAAAAA 350  
 35 ATGGACGTTT GAACCAACTG TGGCAGTTGG TGAAGAAGTG TCGGCAGGTG 400  
 ACATCGTTCG TGTGTTTCAA GAAACACCGA TTATTCAACA TAAAATTATG 450  
 GTGCCTTTCG GCGTTTCAGG AACGATTGCC GAAATTAAAG CAGGTGACTT 500  
 TGCCATTGAT GAAACAGTTT ACTCAGTGGG AACGGCTAAA GGAACGGAAA 550  
 GTTTTAGCAT GATGCAAAAA TGGCCCGTTC GGCGGGGACG TCCCATTTTA 600  
 40 GAAAAACTAA GTCCCAAAGT ACCGATGGTG ACCGGACAAC GCGTAATTGA 650  
 TACCTTTTTT CCAATTACGA AAGGCGGAGC GGCAGCAGTT CCAGGACCAT 700  
 TTGGCGCTGG AAAAACAGTC GTTCAGCACC AAATTGCTAA GTGGGCCGAT 750  
 GTCGACTTAG TCGTTTACGT TGGTTGTGGG GAACGCGGGA ATGAAATGAC 800  
 AGATGTTTTA AATGAATTTT CAGAATTAAT TGACCCAACA ACTGGTGAGT 850  
 45 CTTTGATGAA TCGGACGATT TTAATTGCGA ATACGTCAA TATGCCGGTA 900  
 GCGGCACGGG AAGCCTCGAT TTATACAGGG ATTACCATTG CAGAATATTT 950  
 CCGTGATATG GGTTA 965

50

## 2) INFORMATION FOR SEQ ID NO: 713

## (i) SEQUENCE CHARACTERISTICS:

55 (A) LENGTH: 1737 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

60

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Methanosarcina barkeri*

(C) ACCESSION NUMBER: extracted from J04836

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 713

```

GTGGAAGTAA AAGGTGAAAT TTATCGTGTG TCTGGGCCTG TCGTCACCGC      50
CATCGGCTTG CAGGCAAAAA TGTATGACCT GGTCAAAGTC GGTAATGAAG      100
GTTTAATGGG TGAAGTCATT CAGATATTAG GGCCCAAGAC CATCATCCAG      150
10 GTATATGAAG AGACCGCAGG TATCAAGCCA GGGGAACCCT GTGTATCTAC      200
AGGGTCGTCT CTGTCCGTAG AACTTGGTCC GGGTCTTCTT TCCAGTATTT      250
ATGACGGGGT TCAAAGGCCCT CTGCACGTCC TGCTTGAAAA AATGGGTTAGC      300
TTCATCCAGA GAGGTGTCAG CGCAGATGGG CTTGATCATA AGAAACTCTG      350
GGATTTCAAA CCCATTGTCA AGAAGGGCGA TTCCGTAAAA GGTGGAGACG      400
15 TAATTGGTGT TGTACAGGAA ACCGTGAATA TTGAACATAA GATCATGGTG      450
CCTCCTGATA TCTCAGGTAC AATTTCCGAC ATAAAGAGCG GAAACTTTAC      500
GGTAGTAGAC ACAATCTGTA CTCTGACTGA TGGGACCGAA TTGCAGATGA      550
TGCAGAGGTG GCCTGTTTGA AGACCCAGAC CTGTGAAGGC AAAACTTACT      600
CCAACCAGGC CTCTGGTTAC AGGAATGAGA ATCCTTGATG GGCTTTTCCC      650
20 TGTGGCAAAA GGCGGAACAG CTGCAATCCC CGGACCTTTC GGATCGGGAA      700
AGACCGTAAC TCAGTAGTCG CTTGCAAAAT GGAGTGATAC CGAAATTGTG      750
GTCTACATCG GTTGTGGTGA GCGTGGAAAC GAAATGGCAG ATGTTCTGAG      800
CGAATTCCCT GAACTCGAAG ATCCGCAGAC CGGGCGCCCA CTTATGGAGC      850
GTACTGTTCT TATCGCTAAC ACTTCAAACA TGCCTGTGGC CGCAAGAGAA      900
25 GCATCTGTGT ATACCGGAAT CACCATTGCA GAATACTACC GTGACATGGG      950
ATTAGATGTA TCCCTTATGG CAGACTCCAC CTCAAGGTGG GCAGAAGCCA      1000
TGAGAGAAAT CTCTTCCCGT CTGGAAGAAA TGCCTGGTGA AGAAGGTTAC      1050
CCAGCATACC TGTCTGCAAG ACTGGCCGAA TTCTACGAGC GTGCCGGGGT      1100
TGCGGAGAGT CTTTGCGGGC AAACAGGTTT CATTACTGTT ATTGGAGCAG      1150
30 TATCTCCACC TGGCGGTGAC TTCTCAGAGC CTGTTACACA GAATACCCTG      1200
CGTATCGTAA AAGTGTTCTG GGCTCTCGAT GCCAAACTAT CTCAGAGGCG      1250
TCACTTCCCG GCCATCAACT GGCTGAACAG TTACAGTCTG TATAAGGACA      1300
GTCTTAATGA CTGGTTTGCA GATAATGTGG CTCCTGATTA TGTGCCTTTG      1350
AGGGAAAGAG CAATGGAAAT GCTCCAGACA GAATCTGAAC TGCAGGAAAT      1400
35 CGTGCAGCTT GTAGGTTCCG ATGCTCTGCC AGACGACCAG CAGCTTCTGC      1450
TTGAAATCAC CCGTATGCTT AGGGAAATTT TCCTGCAGCA GAATGCATTC      1500
CACCCAGTAG ATGCATACAG CCCGTTTCGAT CAGCAGTACA AGATCCTTAA      1550
GGCAATCATG AAATGGGGAG ACGCTGCGAT GGATGCCTTG AAATCAGGTG      1600
TTCCCGTAAC TGAAATTATC AAGCTTGAAT CCAAAAATGT GCTTGCTAAG      1650
40 GTCAAGTACG AAGAGAAGTT TGATGAGTCT ATGAATGCTG TCCTGGCACA      1700
GATGGATAAA GAGTTTGCAT CCCTGAGAGG TAGGTAA      1737

```

## 45 2) INFORMATION FOR SEQ ID NO: 714

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1785 bases

(B) TYPE: Nucleic acid

50 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 55 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Methanosarcina jannaschii*

(C) ACCESSION NUMBER: extracted from U67477

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 714

60



	ATGAAAAGAG	AGGTTGAGAA	TATGCCAGTT	GTTGGTAAGA	TTATTAAAAAT	50
	CGCAGGGCCT	GTTGTAGTTG	CAGAGGGAAT	GAAAGGAGCT	CAGATGTATG	100
	AGGTCGTTAA	AGTAGGAGAA	GAGAAATTGA	CTGGAGAAAT	CATTCAGTTG	150
	CACGATGATA	AAGCAGTTAT	TCAGGTTTAT	GAAGAAACAT	CTGGAATTAA	200
5	ACCAGGAGAG	CCAGTTGTTG	GTACTGGAGC	TCCATTGTCT	GTTGAATTAG	250
	GGCCAGGGAT	GTTAAGAGCT	ATGTATGATG	GTATTCAGAG	GCCTTTAACA	300
	GCAATTGAAG	AGAAAACAGG	TTCAATCTTT	ATCCCAAGAG	GAGTTGATGT	350
	CCCTGCATTA	CCAAGAGATA	TAAAATGGGA	ATTTAAACCA	GTGGTAAATG	400
	AAGGAGATTA	TGTTGAAGAA	GGAGACATAA	TTGGAACGTG	TGATGAAACT	450
10	CCTTCAATAG	TTCATAAAAT	CTTAGTTCCA	ATTGGTGTTA	AAGGAAAAAT	500
	TGTTGAAATA	AAAGAGGGTA	AATTTACAGT	TGAAGAGACA	GTTGCAGTTG	550
	TAGAAACAGA	AAATGGAGAA	AGGAAAGAAA	TTACAATGAT	GCAAAAATGG	600
	CCAGTAAGAA	AACCAAGACC	ATATAAAGAG	AACTACCTC	CAGAAATTCC	650
	ATTAATTACA	GGGCAAAGAG	TTGAAGACAC	TTTCTTTACA	TTAGCAAAAG	700
15	GAGGAACAGC	AGCAATTCCA	GGTCCATTCT	GTTTCAGGAA	AACGGTTACT	750
	CAGCATCAGT	TGGCAAAGTG	GTCTGACGCT	GATGTCGTTG	TTTATATCGG	800
	ATGTGGAGAA	AGAGGAAACG	AGATGACAGA	GGTTATTGAA	GAGTTCCCAC	850
	ACTTAGAAGA	TATTAGAACT	GGAAACAAAT	TAATGGATAG	AACTGTATTA	900
	ATAGCCAACA	CATCAACAT	GCCTGTCGCT	GCAAGGGAAG	CATCTGTCTA	950
20	TACAGGAATT	ACAATTGCAG	AGTACTTCAG	AGATATGGGT	TATGGAGTTT	1000
	TATTAACAGC	AGATTCAACA	TCAAGATGGG	CAGAGGCAAT	GAGAGAAATT	1050
	TCAGGTAGAT	TGGAAGAAAT	GCCAGGGGAA	GAAGGGTATC	CAGCATACTT	1100
	AGCTTCAAGA	TTGGCTCAGT	TCTATGAAAG	AGCTGGAAGA	GTTATAACCT	1150
	TAGGGAAAGA	TAACAGACAA	GGATTCGTTT	GTATCGTTGG	AGCTGTTTCA	1200
25	CCACCAGGAG	GGGACTTCTC	AGAACCAGTT	ACATCAAACA	CACTAAGGAT	1250
	AGTTAAGGTA	TTCTGGGCGT	TAGATGCAAA	CTTGGCAAGA	AGAAGACACT	1300
	TCCCAGCTAT	CAACTGGTTG	CAGAGTTATT	CATTATACAT	TGATGATGTT	1350
	ACAGAGTGGT	GGAACACAAA	TACTGGTCCA	GATTGGAGAC	AATTAAGAGA	1400
	TGAAGCAATG	AGCTTATTAC	AAAAAGAGGC	AGAGTTGCAA	GAGATTGTTC	1450
30	AGTTAGTTGG	GCCTGATGCA	TTGCCAGATA	GGGAGAGAGT	TATTTTAGAA	1500
	GTTGCAAGAA	TGTTGAGGGA	GGATTTCTTA	CAGCAAGATG	CGTTTGATGA	1550
	GGTAGATACC	TACTGTCCCTC	CAATGAAACA	GTACTTAATG	TTAAAGATAA	1600
	TTATGACATT	CTACCAAGAA	GCATTGAAGG	CAGTTGAAAG	AGGAGTTGAA	1650
	CCAGCTAAGA	TTTTAGGAGT	TTCAGTTAAG	CAAGATATTG	CAAGAATGAA	1700
35	ATACATCCCA	CACGATGAGT	TTATAAATGT	TAAATCAAAA	GAAATAATGG	1750
	AGAAAATTAA	GAATGAATTA	GGTTCATTAA	ACTAA		1785

## 40 2) INFORMATION FOR SEQ ID NO: 715

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1354 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 50 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*  
(B) STRAIN: W83

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 715

55	GTAGCGAGGT	GATCAAGGTG	ATCGGCAAAA	ATGCTTATGT	GCAGGTTTTTC	50
	GAAAGTACTC	GCGGTATGCA	CGTAGGAGAT	GAGGCAGAGT	TTACCGGCAG	100
	TATGCTTGAG	GTAACGCTCG	GCCCCGGTAT	GCTTTTGAAG	AACTACGACG	150
	GTCTGCAACA	CGACTTGGAC	AAGATGGACG	GGATCTTCCT	CAAACGAGGC	200
60	GATTATACTC	CCGCTCTCGA	TGACGACAAG	CTGTGGGACT	TCAAGCCTTT	250

	GGCCAATGTG	AACGACAATG	TGATCGCAGG	CTCATGGCTC	GGAGAGGTGA	300
	CGGAAAATTT	CCAACCGCAC	AAGATCATGG	TACCTTTCGT	TTTCGAAGGC	350
	AATTACAAGG	TGAAGAGTCT	GGCCAAAGCC	GGTTCGTACA	AAGTGAACGA	400
	TGTGATCGCT	GTGGTAACGG	ATCAGGACGG	GAAAGACCAC	AATGTAACCA	450
5	TGGTGCAGAA	ATGGCCGGTG	AAACGTGCTA	TCACTTGCTA	TCGCGAGAAG	500
	CCGCGTCCTT	TCAAACGTGT	CGAAACGGGT	ATCCGTATCA	TCGACACTTT	550
	CAACCCCATC	GTAGAGGGTG	GTACGGGATT	TATCCCCGGT	CCTTTCGGTA	600
	CGGGAAAGAC	GGTGCTCCAG	CATGCTATCT	CGAAGCAGGC	GGAAGCCGAT	650
	ATCGTGATCA	TTGCAGCCTG	TGGCGAGCGT	GCAAACGAGG	TTGTGGAGAT	700
10	CTTTGCGGAA	TTCCCCCACC	TGAATGACCC	CCACACGGGA	CGTAAATTGA	750
	TGGAACGTAC	CATTATTATT	GCTAATACGT	CGAATATGCC	TGTGGCTTCG	800
	CGTGAGGCAT	CCGTATATAC	GGCCATGACG	ATAGCCGAGT	ACTATCGCTC	850
	CATGGGCCCTT	CGCGTGCTGA	TGATGGCAGA	CTCCACTTCG	CGTTGGGCAC	900
	AGGCTCTCGG	TGAGATGTCT	AACCGTCTGG	AAGAGCTTCC	CGGACCGGAT	950
15	GCTTTTCCGA	TGGACTTGTC	AGCTATCGTA	GCCAACTTCT	ACGCTCGTGC	1000
	AGGATACGTT	TACCTGAACA	ACGGTTTCGGC	CGGTTCGGTA	ACG TTCATCG	1050
	GTACGGTATC	TCCCGCCGGT	GGTAACCTCA	AAGAGCCTGT	GACGGAAAAC	1100
	ACCAAGAAAG	TGGCTCGCTG	CTTCTATGCT	TTGGAGCAGA	ATCGTGCCGA	1150
	CCGCAAACGT	TATCCGGCTG	TAAACCCCAT	CGATAGTTAC	TCGAAGTACA	1200
20	TCGAATATCC	CGAATTCGAG	AGCTATATAT	CGAACCACAT	CAGTTTACTC	1250
	ATTATATTCA	ACAGAAAGTTG	GCAGATATAT	GGATCAAATC	TTACAGCAGA	1300
	CTGGACTACT	AAGGTGAATG	AGCTGAAGAT	GCGCTTGCAT	CAGGGTAAAG	1350
	AAAT					1354

25

## 2) INFORMATION FOR SEQ ID NO: 716

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1788 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: Type 4
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 716

	TTTGACTCAA	GGGAAGATTA	TAAAGTATC	GGGACCTCTA	GTTATTGCAT	50
	CAGGTATGCA	GGAGGCTAAT	ATTCAAGATA	TTTGCCGTGT	AGGTAAGCTA	100
45	GGGTTAATCG	GTGAAATTAT	TGAAATGAGA	AGAGATCAGG	CATCTATCCA	150
	AGTCTATGAA	GAAACATCTG	GTCTTGGTCC	GGGAGAACCT	GTTGTTACAA	200
	CTGGAGAACC	TCTCTCGGTT	GAATTAGGGC	CAGGATTGAT	TTCTCAAATG	250
	TTTGATGGCA	TACAACGCCC	ATTAGATCGA	TTTAAATTGG	CTACTCATAA	300
	TGATTTTCTA	GTTTCGTGGGG	TAGAAGTTCC	AAGTTTGGAT	AGAGATATTA	350
50	AGTGGCATTT	TGATTCCACT	ATAGCAATTG	GTCAAAAAGT	GAGTACGGGT	400
	GATATTCTTG	GAAGTGTCAA	GGAAACCGAG	GTAGTTAATC	ATAAAATTAT	450
	GGTTCCTTAT	GGAGTATCTG	GAGAAGTCGT	TTCTATTGCA	TCTGGCGATT	500
	TTACAATTGA	TGAAGTTGTA	TATGAAATAA	AAAAATTGGA	CGGTAGTTTC	550
	TATAAAGGAA	CGCTTATGCA	AAAATGGCCT	GTCCGCAAGG	CGCGTCCTGT	600
55	TTCTAAACGT	TTAATTCCAG	AAGAACCATT	AATCACAGGT	CAACGAGTTA	650
	TTGATGCATT	CTTTCCAGTA	ACCAAAGGGG	GAGCTGCAGC	AGTTCCTGGA	700
	CCGTTTGGAG	CAGGAAAGAC	AGTTGTACAA	CACCAAGTAG	CTAAATTTGC	750
	CAATGTTGAT	ATTGTTATTT	ATGTCGGTTG	TGGAGAACGT	GGAAATGAAA	800
	TGACGGATGT	ACTGAATGAG	TTTCCTGAGT	TGATTGACCC	TAATACCGGA	850
60	CAATCAATTA	TGCAACGGAC	AGTTCTGATT	GCTAATACTT	CAAATATGCC	900

TGTTCGCTGCT CGTGAGGCTT CAATTTATAC AGGAATTACC ATGGCTGAGT 950  
 ATTTTCGTGA TATGGGCTAC TCTGTCGCCA TTATGGCTGA TTCAACTTCA 1000  
 CGTTGGGCAG AAGCGCTACG TGAAATGTCA GGACGTCTAG AAGAAATGCC 1050  
 TGGTGATGAG GGTATCCTG CTTATCTGGG AAGTCGTATC GCTGAATATT 1100  
 5 ATGAAAGAGC AGGACGTTCT CAGGTTCTAG GGCTTCCAGA ACGTGAAGGA 1150  
 ACGATTACTG CTATTGGAGC TGTATCGCCA CCTGGTGGAG ATATTTTCAGA 1200  
 ACCAGTTACT CAAAACACTT TACGGATTGT GAAAGTTTTT TGGGGGCTTG 1250  
 ATGCTCCGTT GGCACAGCGA CGTCATTTTC CTGCAATTAA CTGGCTTACA 1300  
 TCTTATTCAC TATATAAAGA CAGTGTGGGC ACTTATATAG ATGGTAAAGA 1350  
 10 GAAGACAGAT TGGAAATAGTA AAATAACTCG TGCGATGAAC TACTTACAAC 1400  
 GGAATCTAG TTTAGAGGAA ATTGTTTCGTC TTGTTGGAAT TGATTCTCTG 1450  
 TCTGATAATG AACGACTAAC GATGGAAATT GCTAAACAAA TTCGAGAAGA 1500  
 TTATTTGCAA CAGAACGCTT TTGATTCGGT AGATACATTC ACTTCGTTTG 1550  
 CAAAACAAGA AGCAATGCTA AGTAATATTC TCACTTTTGC TGATCAGGCA 1600  
 15 AATCATGCTT TAGAGTTGGG TTCTTACTTT ACAGAGATTA TGGAAGGTAC 1650  
 CGTGGCAGTT CGAGACCGTA TGGCGAGAAG TAAATATGTT TCAGAAGATA 1700  
 GATTAGATGA AATCAAAAT ATATCAAATG AGATTACACA TCAAATTCAT 1750  
 TTGATATTAG AAACAGGAGG TCTATAAATG AGTGTAT 1788

20

## 2) INFORMATION FOR SEQ ID NO: 717

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 823 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
 30 (ii) MOLECULE TYPE: Genomic DNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Burkholderia mallei*  
 (B) STRAIN: GB8

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 717

CGATCCTGGT GTGCTCGGCC GCTGACGGCC CGATGCCGCA AACGCGTGAG 50  
 CACATCCTGC TGGCGCGTCA GGTGCGGTGTG CCGTACATCA TCGTGTTCCT 100  
 40 GAACAAGTGC GACATGGTGG ACGACGCGGA GCTGCTCGAG CTGGTCGAAA 150  
 TGGAAGTGCG CGAACTGCTG TCGAAGTACG ACTTCCCGGG CGACGACACG 200  
 CCGATCATCA AGGGTTCGGC GAAGCTGGCG CTGGAAGGCG ACAAGGGCGA 250  
 GCTGGGCGAA GTGGCGATCA TGAACCTGGC CGACGCGCTG GACACGTACA 300  
 TCCCGACGCC GGAGCGTGCG GTCGACGGCG CGTTCCTGAT GCCGGTGGAA 350  
 45 GACGTGTTCT CGATCTCGGG CCGTGGTACG GTGGTGACGG GTCGTGTCGA 400  
 GCGCGGCGTG ATCAAGGTTG GCGAGGAAAT CGAAATCGTC GGTATCAAGG 450  
 CGACGGCGAA GACGACCTGC ACGGGCGTGG AAATGTTCCG CAAGCTGCTG 500  
 GACCAGGGTC AGGCGGGCGA CAACGTCGGT ATCCTGCTGC GCGGCACAAA 550  
 GCGTGAAGAC GTGGAGCGCG GCCAGGTTCT GGCGAAGCCG GGTTCGATCA 600  
 50 CGCCGCACAC GCACTTCACG GCAGAAGTGT ACGTGCTGAG CAAGGACGAA 650  
 GGCGGCCGCC ACACGCCGTT CTTCAACAAC TACCGTCCGC AGTTCTACTT 700  
 CCGTACGACG GACGTGACGG GCTCGATCGA GCTGCCGAAG GACAAGGAAA 750  
 TGGTGATGCC GGGCGACAAC GTGTCGATCA CCGTGAAGCT GATCGCGCCG 800  
 ATCGCGATGG AAGAAGGTCT GCG 823

55

## 2) INFORMATION FOR SEQ ID NO: 718

- 60 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 824 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Burkholderia pseudomallei*  
 (B) STRAIN: 1026B

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 718

15	GCGATCCTGG	TGTGCTCGGC	CGCTGACGGC	CCGATGCCGC	AAACGCGTGA	50
	GCACATCCTG	CTGGCGCGTC	AGGTCGGTGT	GCCGTACATC	ATCGTGTTCC	100
	TGAACAAGTG	CGACATGGTG	GACGACGCGG	AGCTGCTCGA	GCTGGTCGAA	150
	ATGGAAGTGC	GCGAACTGCT	GTCGAAGTAC	GACTTCCCAG	GCGACGACAC	200
	GCCGATCATC	AAGGGTTCGG	CGAAGCTGGC	GCTGGAAGGC	GACAAGGGCG	250
	AGCTGGGCGA	AGTGGCGATC	ATGAACCTGG	CCGACGCGCT	GGACACGTAC	300
20	ATCCCGACGC	CGGAGCGTGC	GGTCGATGGC	GCGTTCCTGA	TGCCGGTGGA	350
	AGACGTGTTT	TCGATCTCGG	GCCGTGGTAC	GGTGGTGACG	GGTCGTGTCG	400
	AGCGCGGCGT	GATCAAGGTT	GGCGAGGAAA	TCGAAATCGT	CGGTATCAAG	450
	GCGACGGCGA	AGACGACCTG	CACGGGCGTG	GAAATGTTCC	GCAAGCTGCT	500
	GGATCAGGGT	CAGGCGGGCG	ACAACGTCGG	TATCCTGCTG	CGCGGCACGA	550
25	AGCGTGAAGA	CGTGGAGCGC	GGCCAGGTTT	TGGCGAAGCC	GGGTTTCGATC	600
	ACGCCGCACA	CGCACTTCAC	GGCTGAAGTG	TACGTGCTGA	GCAAGGACGA	650
	AGGCGGCCGC	CACACGCCGT	TCTTCAACAA	CTACCGTCCG	CAGTTCTACT	700
	TCCGTACGAC	GGACGTGACG	GGCTCGATCG	AGCTGCCGAA	GGACAAGGAA	750
	ATGGTGATGC	CGGGCGACAA	CGTGTGATC	ACGGTGAAGC	TGATCGCGCC	800
30	GATCGCGATG	GAAGAAGGTC	TGCG			824

2) INFORMATION FOR SEQ ID NO: 719

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 800 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium beijerincki* (deposited as  
*Clostridium butyricum*)  
 (B) STRAIN: ATCC 8260

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 719

50

	TGTATCAGCA	GCAGATGGTC	CAATGCCACA	AACAAGAGAA	CATATACTAC	50
	TAGGATCAAG	AGTTGGTATC	CAATATATCG	TAGTATTCTT	AAATAAAGCA	100
	GATATGGTAG	ACGATCCAGA	ATTATTAGAA	TTAGTAGAAA	TGGAAGTAAG	150
	AGAATTATTA	AGCGAATATG	ACTTCCCAGG	AGACGATATT	CCAGTAATAA	200
55	CAGGATCAGC	ATTAAAAGCA	TTAGAAAATC	CAACAGATGA	AGAAGCAATT	250
	AAGCCAATCA	TGGATTTAAT	GGAAGCAGTA	GATAGCTATA	TCCCAACTCC	300
	AGAAAGAGCA	ACAGATAAGC	CATTCTTAAT	GCCAATCGAA	GATGTATTCA	350
	CAATTACAGG	AAGAGGAACA	GTTGCAACAG	GAAGAGTTGA	AGCTGGAGTA	400
	CTTCATGTAG	GAGATGAAGT	AGAAATCGTT	GGATTAACAG	AAGAAAAGAA	450
60	GAAAGTTGTA	GTAAGTGGAA	TCGAAATGTT	CAGAAAGTTA	TTGGATGAAG	500

```

5  CACAAGCTGG AGATAACATC GGAGCATTAT TAAGAGGAGT TCAAAGAACT 550
   GATATTGAAA GAGGTCAAGT TTTATCAAAA CCAAATTCAG TACACCCTCA 600
   CACTAAATTT GTAGGTCAAG TATACGTACT TAAAAAAGAA GAAGGTGGAA 650
   GACATACTCC ATTCTTTGAT GGATACAGAC CACAATTCTA TTTCAGAACA 700
   ACAGACGTTA CAGGRTCAAT CAAGTTACCA GATGGAATGG AAATGGTAAT 750
   GCCTGGAGAT CACATTGATA TGAATGTTGA ATTAATCACT CCAATCGCAA 800

```

10 2) INFORMATION FOR SEQ ID NO: 720

(i) SEQUENCE CHARACTERISTICS:

```

15  (A) LENGTH: 799 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Clostridium innocuum
    (B) STRAIN: ATCC 14501

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 720

```

25  GGTGCTATCC TGGTTGTTGC TGCATCTGAT GGTCCCTATGC CTCAGACTCG 50
    TGAGCACATC CTGCTTGCTC GTCAGGTAGG TGTTCCCTTAC ATCGTTGTAT 100
    TCCTGAACAA ATGCGACATG GTTGATGACG AAGAACTGAT CGACCTTGTT 150
    GAAATGGAAG TACGTGAGCT GTTAAGCGAG TACGGATTCG ACGGAGATAA 200
30  CGCTCCGGTT ATCCGTGGTT CTGCACTGAA GGCTCTGGAA GGTGACGACA 250
    AATACGTTGG CGCTATCAAA GAACGTATGG ATGCAGTTGA TGAATTCATC 300
    CCAGATCCAA CTCGTGAAAC TGACAAACCA TTCCTGATGT CTGTAGAAGA 350
    CGTTATGACA ATCACAGGAC GTGGTACAGT TGCTACAGGA CGTGTTGAGC 400
    GTGGGGTAGT AAAACTGGGA GAAGAAGTTG AAATCGTTGG TATCAAGGAT 450
35  ACTCAGAAAA CTGTTGTTAC CGGACTGGAA ATGTTCCGTA AGCAGCTGGA 500
    CTTTCGAGAA TCCGGAGACA ACATCGGTGC TCTGCTGCGT GGTATCAACC 550
    GTGACCAGAT TCAGCGTGGA CAGGTTCTTG CTAAACCAGG ATCCGTACAT 600
    CCACACACAA AGTTCAAGGC TCAGGTTTAT GTATTAACAA AAGAAGAAGG 650
    TGGACGTCAC ACTCCATTGC TTTCTAATC CCGTCCTCAG TTCTACTTCC 700
40  GTACAACTGA CGTAACTGGT GTTATTACAT TACCGGAAGG AACTGAAATG 750
    GTTATGCCTG GTGACAACGT TGAAATGAAC GTTGAGCTGA TTGCTCCAA 799

```

45 2) INFORMATION FOR SEQ ID NO: 721

(i) SEQUENCE CHARACTERISTICS:

```

50  (A) LENGTH: 789 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

(ii) MOLECULE TYPE: Genomic DNA

55 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Clostridium novyi
    (B) STRAIN: ATCC 19402

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 721

60

	TTGTATCAGC	AGCAGATGGT	CCAATGCCAC	AAACAAGAGA	ACACATCCTA	50
	TTAGCATCAA	GAGTAGGAGT	TAACCACATA	GTAGTATTCT	TAAACAAAGC	100
	AGACCAAGTA	GATGATCCAG	AATTACTAGA	ATTAGTAGAA	ATGGAAGTAA	150
	GAGAATTATT	AAGCGAATAC	GGATTTGACG	GAGACGAATG	TCCAGTAGTA	200
5	GTAGGATCAG	CATTAAAAGC	AATCGAAGAA	GGGGATGACC	AATGCATCCT	250
	AGACTTAATG	AAAGCTGTAG	ATGAATATAT	CCCAACTCCA	GAAAGAGCAA	300
	CAGATCAACC	ATTCTTAATG	CCTGTAGAAG	ATGTATTTAC	AATTACAGGA	350
	AGAGGAACAG	TTGCAACAGG	AAGAGTTGAA	AGAGGAGTAC	TACACGTAGG	400
	AGATGAAGTA	CAAATCGTAG	GAATGAAAAG	AGAAATCGGA	AAGACAACAA	450
10	TCACAGGAGT	AGAAATGTTT	AGAAAGATGT	TAGATGAAGC	AATGGCTGGA	500
	GATAACATCG	GAGCATTATT	AAGAGGAGTA	CAAAGAGACG	AAATCGAAAG	550
	AGGTCAAGTA	CTAGCAAAAC	CAGGTTTCAGT	AACACCTCAC	AAAAAATTCG	600
	TAGGTCAAGT	TTACGTATTA	AAGAAAGAAG	AAGGTGGAAG	ACACACTCCA	650
	TTCTTTAACG	GATACAGACC	ACAATTCTAC	TTCAGAACAA	CAGACGTAAC	700
15	AGGATCAATC	GCTTTACCAG	AAGGAGTAGA	AATGGTAATG	CCAGGAGACC	750
	ATATAGACAT	GAACGTAGAA	TTAATCACAC	CAGTAGCAA		789

## 20 2) INFORMATION FOR SEQ ID NO: 722

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 bases  
 (B) TYPE: Nucleic acid  
 25 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium septicum*  
 (B) STRAIN: ATCC 12464

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 722

35	GTTCAGCAGC	AGACGGTCCA	ATGCCACAAA	CAAGAGAACA	TATACTACTA	50
	GCATCAAGAG	TTGGTGTGTA	CTATATCGTA	GTATTCTTAA	ACAAGGCAGA	100
	TATGGTAGAT	GACGAAGAAT	TATTAGAATT	AGTAGAAATG	GAAGTTAGAG	150
	AATTATTATC	AGAATACAAC	TTCCCAGGAG	ATGATATTCC	AGTAATCAAG	200
40	GGATCAGCTT	TAGTAGCATT	AGAAAACCCA	ACAGATGAAA	AATCAATCGC	250
	TCCAATCTTA	GAATTAATGG	AAGCAGTAGA	TAGCTACATT	CCAACACCAG	300
	AAAGAGCAAC	AGATAAGCCA	TTCTTAATGC	CAGTAGAAGA	TGTATTACAC	350
	ATAACTGGTA	GAGGAACAGT	TGCAACAGGA	AGAGTTGAAA	GAGGAGTTCT	400
	TCATGTAGGA	GACGAAGTAG	AAATCGTTGG	ATTATCAGAA	GAAAGCAGAA	450
45	AAGTAGTAGT	AACAGGAATA	GAAATGTTCA	GAAAGTTACT	AGACGAAGCA	500
	CAAGCTGGAG	ATAATGTTGG	AGTACTTTTA	AGAGGTGTTC	AAAGAACAGA	550
	TATCGAAAGA	GGTCAAGTAT	TAGCAAAGAC	TGGATCAGTT	AAGCCACACA	600
	GCAAGTTCGT	AGGTCAAGTA	TACGTACTTA	AGAAAGAAGA	AGGTGGAAGA	650
	CATACTCCAT	TCTTCGATGG	ATACAGACCA	CAATTCTACT	TCAGAACAAC	700
50	AGACGTTACT	GGATCAATCA	AATTACCAGA	CGGAATGGAA	ATGGTTATGC	750
	CAGGAGACCA	CATTGATATG	AACGTTGAAT	TAATCACTCA	AGTAGCAA	798

## 55 2) INFORMATION FOR SEQ ID NO: 723

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 799 bases  
 (B) TYPE: Nucleic acid  
 60 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium tertium*  
 (B) STRAIN: ATCC 14573

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 723

```

10 GTTCAGCAGC AGATGGTCCA ATGCCTCAAA CAAGAGAACA CATACTACTA      50
   GCTTCAAGAG TTGGTGTTGA CTACATAGTT GTTTTCTTAA ACAAGGCAGA      100
   TATGGTAGAT GACGAAGAAT TATTAGAATT AGTTGAAATG GAAGTAAGAG      150
   AATTATTATC AGAATACAAC TTCCCAGGAG ATGATATTCC AGTAATAAAG      200
15 GGTTCAGCTT TACAAGCATT AGAAAACCCA ACAGATGAAA AAGCAATCGC      250
   TCCAATCCTT GAGTTAATGG AAGCTGTAGA TAGCTACATT CCAACTCCAG      300
   AAAGAGCAAC AGATAAGCCA TTCTTAATGC CAGTAGAAGA TGTATTACAA      350
   ATCACTGGTA GAGGAACAGT TGCTACAGGA AGAGTTGAAA GAGGAGTTCT      400
   TCACGTAGGA GACGAAGTAG AAATCGTTGG ATTATCAGAA GACAGCAGAA      450
20 AAGTAGTAGT AACAGGAATA GAAATGTTCA GAAAGTTACT AGACGAAGCG      500
   CAAGCTGGAG ACAACGTAGG AGTTCTTTTA AGAGGAGTTC AAAGAAGTGA      550
   CATCGAAAGA GGTCAAGTTT TAGCAAAAGT TGGATCAGTT AAGCCACACA      600
   AGAAATTTGT AGGTCAAGTA TACGTACTTA AAAAAGAAGA AGGTGGAAGA      650
   CATACTCCAT TCTTCGATGG ATACAGACCA CAATTCTACT TCAGAACAAC      700
25 AGAYGTTACT GGTTCATCA AGTTACCAGA TGGAAATGGAA ATGGTTATGC      750
   CAGGAGACCA CATTGATATG AACGTTGAAT TAATCACTCA AGTAGCTAT      799

```

30 2) INFORMATION FOR SEQ ID NO: 724

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 801 bases  
 (B) TYPE: Nucleic acid  
 35 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

40 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium tetani*  
 (B) STRAIN: ATCC 19406

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 724

```

45 TAGTAAGTGC AGCAGATGGT CCAATGCCAC AAACAAGAGA ACACATACTA      50
   TTAGCATCCA GAGTTGGAGT TGAGCACATA GTAGTATTCT TAAATAAAGC      100
   AGACCAAGTA GATGACGCAG AGTTAATCGA ATTAGTAGAA ATGGAAGTAA      150
   GGGAATTAAT GAACGAATAC GGATTCCTCAG GAGATGACGC ACCAGTAGTA      200
50 GTAGGATCCG CATTAAAAGC ATTAGAAAAT CCAGAAGATG ATGCAGCAAC      250
   ACAATGCATA ATGGACTTAA TGGCAGCAGT AGATGAATAT ATACCAACAC      300
   CAGAAAGAGC AACAGATAAG CCATTCTTAA TGCCAGTAGA AGATATCTTC      350
   ACAATCACAG GAAGAGGAAC AGTTGCAACA GGAAGAGTAG AAAGAGGAAT      400
   TCTAAAAGTA GGAGACGAAA TAGAAATCGT AGGATTAAGT GATGAAAGCA      450
55 AGAAATCAGT AATCACAGGA ATAGAAATGT TCAGAAAACCT ATTAGATGAA      500
   GCACAAGCAG GAGATAACAT CGGAGCATTA TTAAGAGGTG TTCAAAGAGA      550
   TGAAATCCAA AGAGGTCAAG TATTAGCAGC AACAGGATCA GTAAAACCAC      600
   ATAAGAGTTT TACAGGTCAA GTATATGTAT TAAAGAAAGA AGAAGGAGGA      650
   AGACACACTC CATTCTTTAA CGGATACAGA CCACAATTCT ACTTTAGAAC      700
60 AACAGACGTA ACAGGTTCAT TCGCACTACC AGAAGGAGTA GAAATGGTAA      750

```

TGCCAGGAGA CCACATAGAC ATGAAGGTAG AATTAATAAC AAGAGTAGCA 800  
A 801

5

## 2) INFORMATION FOR SEQ ID NO: 725

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 633 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## 15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus malodoratus*  
(B) STRAIN: ATCC 43197

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 725

TATGCCTCAA ACTCGTGAAC ACATCTTGTT ATCTCGTAAC GTTGGTGTTT 50  
CTTACATCGT TGTATTCTTA AACAAAATGG ATATGGTTGA TGATGAAGAA 100  
TTACTAGAAT TAGTTGAAAT GGAAGTTCGT GACTTATTGT CAGAATACGA 150  
25 CTTCCCAGGC GACGACACTC CAGTTATCGC TGGTTCAGCT TTGAAAGCTT 200  
TAGAAGGCGA TGCTTCATAC GAAGAAAAAA TCTTAGAATT AATGGCTGCT 250  
GTTGATGAAT ATATCCCAAC ACCAGTTCGT GATACTGACA AACCATTTCAT 300  
GATGCCAGTC GAAGATGTAT TCTCAATCAC TGGTCGTGGA ACTGTTGCAA 350  
CTGGTCGTGT TGAACGTGGA CAAGTTCGCG TTGGTGACGA AGTTGAAATC 400  
30 GTTGGTATTG CTGAAGCAAC TGCTAAAACA ACTGTTACAG GTGTTGAAAT 450  
GTTCCGTAAT TTGTTAGATT ACGCTGAAGC AGGCGATAAC ATTGGTGCAT 500  
TGTTACGTGG TGTTCACGTG GAAGACATCC AACGTGGACA AGTATTGGCT 550  
AAACCAGCTT CAATCACTCC ACATACAAAA TTCTCTGCAG AAGTTTACGT 600  
TTTAACTAAA GAAGAAGGCG GACGTCATAC TCC 633

35

## 2) INFORMATION FOR SEQ ID NO: 726

## 40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 623 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
45 (D) TOPOLOGY: Linear

## 45 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Enterococcus sulfureus*  
(B) STRAIN: ATCC 49903

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 726

CACGTGAACA CATCTTGTTA TCTCGTAACG TAGGTGTTCC TTACATCGTT 50  
55 GTATTCTTAA ACAAATGGA TATGGTTGAT GACGAAGAAT TATTAGAATT 100  
AGTAGAAATG GAAGTTCGTG ACTTATTATC AGAATACGAT TTCCCAGGCG 150  
ATGACACTCC AGTTGTTGCA GGTCTGCTT TGAAAGCTTT AGAAGGCGAC 200  
GCTTCTTACG AAGAAAAAAT CATGGAATTA ATGGCTGCAG TTGACGAGTA 250  
CATCCCAACT CCAACTCGTG ACACTGACAA ACCATTTCATG ATGCCAGTTG 300  
60 AGGATGTATT CTCAATCACT GGACGTGGTA CTGTTGCTAC AGGTCGTGTT 350



GAACGTGGAC AAGTTCGCGT TGGTGACGTT GTAGATATCG TTGGTATCGC 400  
 TGACGAAACT GCTCAAACAA CTGTAACAGG TGTTGAAATG TTCCGTAAAT 450  
 TATTAGACTA CGCTGAAGCA GGCGATAACA TCGGTGCTTT ATTACGTGGT 500  
 GTTGCTCGTG AAGACATCCA ACGTGGACAA GTTTTAGCTA AACCAGCTTC 550  
 5 AATCACTCCA CATACAAAT TCTCTGCTGA AGTATACGTA TTAAGCAAAG 600  
 AAGAAGGTGG ACGTCATACT CCA 623

10 2) INFORMATION FOR SEQ ID NO: 727

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 646 bases  
 (B) TYPE: Nucleic acid  
 15 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Lactococcus garvieae*  
 (B) STRAIN: ATCC 49156

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 727

25 CGGTCCTATG CCTCAAACCTC GTGAACACAT CTTGTTATCA CGTAACGTTG 50  
 GCGTACCATA CATCGTTGTT TTCTTGAACA AAATGGATAT GGTTGATGAC 100  
 GAAGAATTGC TAGAATTAGT TGAAATGGAA GTTCGTGACC TATTGTCTGA 150  
 ATATGACTTC CCAGGCGACG ATGTTCCCTGT AATCGCTGGT TCTGCTTTGA 200  
 30 AAGCTCTTGA AGGAGATCCT TCATACGAAG AAAAAATCAT GGAATTGATG 250  
 GCTGCAGTTG ACGAATACGT TCCAACCTCCA GAACGTGATA CTGACAAACC 300  
 ATTCATGATG CCAGTCGAAG ACGTATTCTC AATCACTGGA CGTGGTACTG 350  
 TTGCTACAGG CCGTGTTGAA CGTGGACAAG TTCGCGTTGG TGATGAAGTA 400  
 GAAATCGTTG GTATTGCTGA CGAAACTGCT AAAACAACTG TAACAGGTGT 450  
 35 TGAAATGTTT CGTAAATTGT TAGACTATGC TGAAGCAGGG GATAACATTG 500  
 GTGCATTGCT ACGTGGGGTT GCTCGTGAAG ACATCCAACG TGGACAAGTA 550  
 TTGGCTAAAG CTGGTACAAT CACACCTCAT ACAAATTCA AAGCTGAAGT 600  
 TTATGTTTTG ACAAAGAAG AAGGTGGACG TCACACTCCA TTCTTC 646

40 2) INFORMATION FOR SEQ ID NO: 728

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 823 bases  
 (B) TYPE: Nucleic acid  
 45 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

50 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Mycoplasma pirum*  
 (B) STRAIN: ATCC 25960D

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 728

GAGCAATTTT AGTTGTTTCT GCAACTGATG GTCCAATGCC TCAAACCTCGT 50  
 GAACATATCT TATTAGCAGC CCAAGTTGGT GTTCCTAAAA TGGTTGTTTT 100  
 CTTAAACAAA TGTGATGTTG CTTCTGATCC AGAAATGCAA GAATTAGTTG 150  
 60 CTGAAGAAGT AAAAGACTTA TTAAAATCTT ATGGTTTGA TGGTGACAAT 200

```

ACTCCAATTA TTCGTGGTTC AGCATTAGAA GCATTAAATG GTAAACCTGA 250
ATGAGAAGAA AAAATTAAAG AATTAATGAA GGCAGTGGAT GACACTATTC 300
CTGATCCAGT TCGTGATACT GAAAAGCCAT TCTTGTTACC AATTGAAGAC 350
GTAATGACAA TTACAGGTCG TGGTACTGTT GTTACAGGTC GTGTAGAACG 400
5 TGGTACTCTA AAATTAAATG ATGAAGTTGA AATTGTTGGT TTAGGTGAAA 450
CATTTAAATC TGTGTAAACA GGTATTGAAA TGTTCGTAA AGAATTAGAT 500
GAAGCTCGTG CTGGTGACAA TGCTGGTATT TTACTIONCGTG GTGTTGACCG 550
TGGTCAAGTA CAACGTGGTC AAGTTCTTGC TAAACCAGGT TCTATTACTC 600
CTCATACTAA ATTTAAAGCT GAAATTTATG CTTTGAAAAA AGAAGAAGGT 650
10 GGTTCGTCATA CTGCTTTCTT AAACGGTTAT CGTCCTCAAT TCTATTTTCA 700
AACAACGTGAT GTTACAGGTT CTATTAAATT AAAAGATGGA ACTGAAATGG 750
TTATGCCTGG TGACAATACT GAAATCACTG TAGAATTAAT TTCACCAATT 800
GCTTGTAATA AGGAAGTAAG TTT 823

```

15

## 2) INFORMATION FOR SEQ ID NO: 729

```

20 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 826 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

25 (ii) MOLECULE TYPE: Genomic DNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Mycoplasma salivarium
        (B) STRAIN: ATCC 23064

```

30

—(xi) SEQUENCE DESCRIPTION:— SEQ ID NO: 729 —

```

GGAGCAATCT TAGTTGTTGC TGCAACTGAT GGTGCGATGC CTCAAACTCG 50
TGAACACGTT TTACTTGCAA AACAAGTTGG TGTTCTTAAA ATCGTTGTTT 100
35 TCTTAAACAA AATCGATATG TTCAAGCCAG AAGAAAAGAGC CGAAATGGTT 150
GAAATGGTTG AAATGGACAT TCGTGACTTA TTAAACAAAT ACGACTTTGA 200
TGGTGACAAT ACTCCTATTA TTGCTGGTTC AGCATTAAAA GCATTACAAG 250
GTGATCCAGA ATATGAAAAG AATATTCTAG AACTTATGGA TGCAGTAGAC 300
ACATACATTG ATGAACCTAC TCGTGATAAA GATAAACCAT TCTTAATGGC 350
40 TGTAGAAGAC GTATTTACAA TTACTGGTTC TGGAACTGTT GCTACTGGTA 400
GAGTAGAACG TGGTACATTA CATCTAAATG ATGAAGTTGA AATCGTTGGT 450
CTACGTCCAA CAATTAAAC TGTGTACTT GGAATTGAAA TGTTCGTAA 500
AAATTTAAAA GAAGCTCAAG CTGGAGATAA TGCAGGACTT TTACTACGTG 550
GAATTGATAG AGATCAAGTA GAACGTGGAC AAGTTTGGC CAAACCAAAA 600
45 AGTATTATTC CTCACACAGA ATTTGAAGCT GCTGTGTATG TTCTAAAAGC 650
TGAAGAAGGT GGACGTCACA CTCCATTCTT TGAACACTAT AAACCACAAT 700
TTTACTTTTC TACAACCGAC GTTACTGGTG GAATTAAATT CAAACCTGGA 750
CGTGAAATGG TTATGCCTGG CGAAAATGTT GAATTTACAG TTACTTTAAT 800
50 TGCTCCTATT GCAGTTGAAG AAGGAA 826

```

50

## 2) INFORMATION FOR SEQ ID NO: 730

```

55 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 810 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria polysaccharea*

(B) STRAIN: ATCC 43768

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 730

	TGGTATGTTT	CGCAGCYGAY	GGYCCTATGC	CTCAAACCTCG	CGAACACATC	50
10	CTGYTGGCTC	GCCAAGTAGG	YGTACCTTAC	ATCATCGTRT	TCATGAACAA	100
	ATGCGACATG	GTTGACGATG	CCGAGCTGTT	GGAAGTGGTT	GAAATGGAAA	150
	TCCGYGACCT	GYTGTCMAGC	TACGACTTCC	CMGGCGACGA	CTGCCCAATC	200
	GTACAAGGTT	CTGCACTGAA	AGCTTTGGAA	GGYGAYGCTG	GTTACGAAGA	250
	GAAAATCTTC	GAAYTGGCTG	CTGCTTTGGA	CAGCTACATC	CCAACTCCTG	300
15	AGCGTGCTGT	GGACAAACCK	TTYTGTYTGC	CTATCGAAGA	CGTRTTCTCT	350
	ATCTCTGGYC	GYGGTACMGT	AGTAACYGGY	CGTGTAGAGC	GCGGTRTCAT	400
	CCACGTTGGT	GACGAGATYG	AAATCGTAGG	TCTGAAAGAA	ACCCAAAAAA	450
	CCACTTGTAC	CGGCGTTGAA	ATGTTCCGCA	AACTGCTGGA	CGAAGGTCAA	500
	GCWGGTGACA	ACGTAGGCGT	ATTGYTGCGT	GGTACCAARC	GTGAAGACGT	550
20	AGAGCGTGGT	CAAGTATTGG	CYAAACCAGG	YACCATYACT	CCSCACACCA	600
	AATTCAAAGC	AGAAGTATAC	GTACTGAGCA	AAGAAGAGGG	TGGTCGYCAC	650
	ACTCCATTCT	TCGCTAATA	CCGYCCMCAA	TTCTACTTCC	GTACYACTGA	700
	CGTAACYGGT	GCAGTTACTT	TGGAAGAAGG	TGTAGAAATG	GTAATGCCAG	750
	GTGAGAACGT	AACCATTACY	GTAGAACTGA	TTGCGCCTAT	CGCTAGGAAG	800
25	AAGGTCTGCG					810

2) INFORMATION FOR SEQ ID NO: 731

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 813 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*  
serotype Enteritidis

(B) STRAIN: ATCC 13076

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 731

45	CTGGTTGTTG	CTGCGACTGA	CGGCCCGATG	CCGCAGACCC	GTGAGCACAT	50
	CCTGCTGGGT	CGTCAGGTAG	GCGTTCCGTA	CATCATCGTG	TTCCTGAACA	100
	AATGCGACAT	GGTTGATGAC	GAAGAGCTGC	TGGAAGTGGT	TGAAATGGAA	150
	GTTCGYGAAC	TGCTGTCTCA	GTACGACTTC	CCGGGCGACG	ACACTCCGAT	200
50	CGTTTCGTGGT	TCTGCTCTGA	AAGCGCTGGA	AGGCGACGCA	GAGTGGGAAG	250
	CGAAAATCAT	CGAACTGGCT	GGCTTCCTGG	ATTCTTACAT	CCCGGAACCA	300
	GAGCGTGCGA	TTGACAAGCC	GTCCTGCTG	CCGATCGAAG	ACGTATTCTC	350
	CATCTCCGGT	CGTGGTACCG	TTGTTACCGG	TGCTGTAGAA	CGCGGTATCA	400
	TCAAAGTGGG	CGAAGAAGTT	GAAATCGTTG	GTATCAAAGA	GACTCAGAAG	450
55	TCTACCTGTA	CTGGCGTTGA	AATGTTCCGC	AACTGCTGG	ACGAAGGCCG	500
	TGCCGGTGAG	AACGTAGGTG	TTCTGCTGCG	TGGTATCAA	CGTGAAGAAA	550
	TCGAACGTGG	TCAGGTACTG	GCTAAGCCGG	GCACCATCAA	GCCGCACACC	600
	AAGTTCGAAT	CTGAAGTGTA	CATTCTGTCC	AAAGATGAAG	GCGGCCGTCA	650
	CACTCCGTTT	TTCAAAGGCT	ACCGTCCGCA	GTTCTACTTC	CGTACTACTG	700
60	ACGTGACTGG	CACCATCGAA	CTGCCGGAAG	GCGTAGAGAT	GGAATGCCG	750

GGCGACAACA TCAAAATGGT TGTTACCCTG ATCCACCCGA TCGCAATGGA 800  
CGACGGTCTG CGT 813

5

## 2) INFORMATION FOR SEQ ID NO: 732

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 812 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

15

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*  
serotype Gallinarum  
(B) STRAIN: ATCC 9184

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 732

CTGGTTGTTG CTGCGACTGA CGGCCCGATG CCGCAGACCC GTGAGCACAT 50  
CCTGCTGGGT CGTCAGGTAG GCGTTCCGTA CATCATCGTG TTCCTGAACA 100  
25 AATGCGACAT GGTGTGATGAC GAAGAGCTGC TGGAACTGGT TGAAATGGAA 150  
GTTTCGYGAAC TGCTGTCTCA GTACGACTTC CCGGGCGACG ACACTCCGAT 200  
CGTTCGTGGT TCTGCTCTGA AAGCGCTGGA AGGCGACGCA GAGTGGGAAG 250  
CGAAAATCAT CGAACTGGCT GGCTTCTCTG ATTCTTACAT CCCGGAACCA 300  
GAGCGTGCGA TTGACAAGCC GTTCCTGCTG CCGATCGAAG ACGTATTCTC 350  
30 CATCTCCGGT CGTGGTACCG TTGTTACCGG TCGTGTAGAA CGCGGTATCA 400  
TCAAAGTGGG CGAAGAAGTT GAAATCGTTG GTATCAAAGA GACTCAGAAG 450  
TCTACCTGTA CTGGCGTTGA AATGTTCCGC AAAGTCTGCTG ACGAAGGCCG 500  
TGCCGGTGAG AACGTAGGTG TTCTGCTGCG TGGTATCAAA CGTGAAGAAA 550  
TCGAACGTGG TCAGGTACTG GCTAAGCCGG GCACCATCAA GCCGCACACC 600  
35 AAGTTCGAAT CTGAAGTGTA CATTCTGTCC AAAGATGAAG GCGGCCGTCA 650  
CACTCCGTTT TTCAAAGGCT ACCGTCCGCA GTTCTACTTC CGTACTACTG 700  
ACGTGACTGG CACCATCGAA CTGCCGGAAG GCGTAGAGAT GGTAATGCCG 750  
GGCGACAACA TCAAAATGGT TGTTACCCTG ATCCACCCGA TCGCAATGGA 800  
CGACGGTCTG CG 812

40

## 2) INFORMATION FOR SEQ ID NO: 733

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 814 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

50

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*  
serotype Paratyphi B  
(B) STRAIN: ATCC 8759

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 733

60 TGTTGTTGTC TGCGACTGAC GGYCCGATGC CGCAGACCCG TGAGCACATC 50

```

CTGCTGGGTC GTCAGGTAGG CGTTCCGTAC ATCATCGTGT TCCTGAACAA 100
ATGCGACATG GTTGATGACG AAGAGCTGCT GGAAGTGGTT GAAATGGAAG 150
TTCGYGAACT GCTGTCTCAG TACGACTTCC CGGGCGACGA CACTCCGATC 200
GTTTCGTGGTT CTGCTCTGAA AGCGCTGGAA GGCGACGCAG AGTGGGAAGC 250
5 GAAAATCATC GAACTGGCTG GCTTCCTGGA TTCTTACATT CCGGAACCAG 300
AGCGTGCGAT TGACAAGCCG TTCCTGCTGC CGATCGAAGA CGTATTCTCC 350
ATCTCTGGTC GTGGTACCGT TGTTACCGGT CGTGTAAGAAC GCGGTATCAT 400
CAAAGTGGGC GAAGAAGTTG AAATCGTTGG TATCAAAGAG ACTCAGAAGT 450
10 CTACCTGTAC TGGCGTTGAA ATGTTCCGCA AACTGCTGGA CGAAGGCCGT 500
GCTGCTGAGA ACGTAGGTGT TCTGCTGCGT GGTATCAAAC GTGAAGAAAT 550
CGAACGTGGT CAGGTACTGG CTAAGCCGGG CACCATCAAG CCGCACACCA 600
AGTTCGAATC TGAAGTGTAC ATTCTGTCCA AAGATGAAGG CGGCCGTCAT 650
ACTCCGTTCT TCAAAGGCTA CCGTCCGCAG TTCTACTTCC GTACTACTGA 700
CGTGACTGGC ACCATCGAAC TGCCGGAAGG CGTAGAGATG GTAATGCCGG 750
15 GCGACAACAT CAAAATGGTT GTTACCCTGA TCCACCCGAT CGCAATGGAC 800
GACGGTCTGC GTTT 814

```

## 20 2) INFORMATION FOR SEQ ID NO: 734

## (i) SEQUENCE CHARACTERISTICS:

```

25 (A) LENGTH: 828 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

```

----- (A) -- ORGANISM: Salmonella choleraesuis subsp. choleraesuis
          serotype Virchow
          (B) STRAIN: ATCC 51955

```

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 734

```

GGCGCGATCC TGGTTGTTGC TGCGACTGAC GGCCCGATGC CGCAGACCCG 50
TGAGCACATC CTGCTGGGTC GTCAGGTAGG CGTTCCGTAC ATCATCGTGT 100
TCCTGAACAA ATGCGACATG GTTGATGACG AAGAGCTGCT GGAAGTGGTT 150
40 GAAATGGAAG TTCGTGAACT TCTGTCTCAG TACGACTTCC CGGGCGACGA 200
CACTCCGATC GTTCGTGGTT CTGCTCTGAA AGCGCTGGAA GGCGACGCAG 250
AGTGGGAAGC GAAAATCATC GAACTGGCTG GCTTCCTGGA TTCTTACATT 300
CCGGAACCAG AGCGTGCGAT TGACAAGCCG TTCCTGCTGC CGATCGAAGA 350
CGTATTCTCC ATCTCCGGTC GTGGTACCGT TGTTACCGGT CGTGTAAGAAC 400
45 GCGGTATCAT CAAAGTGGGC GAAGAAGTTG AAATCGTTGG TATCAAAGAG 450
ACTCAGAAGT CTACCTGTAC TGGCGTTGAA ATGTTCCGCA AACTGCTGGA 500
CGAAGGCCGT GCCGGTGAGA ACGTAGGTGT TCTGCTGCGT GGTATCAAAC 550
GTGAAGAAAT CGAACGTGGT CAGGTACTGG CTAAGCCGGG CACCATCAAG 600
CCGCACACCA AGTTCGAATC TGAAGTGTAC ATTCTGTCCA AAGATGAAGG 650
50 CGGCCGTCAT ACTCCGTTCT TCAAAGGCTA CCGTCCGCAG TTCTACTTCC 700
GTACTACTGA CGTGACTGGC ACCATCGAAC TGCCGGAAGG CGTAGAGATG 750
GTAATGCCGG GCGACAACAT CAAAATGGTT GTTACCCTGA TCCACCCGAT 800
CGCAATGGAC GACGGTCTGC GTTTCGCA 828

```

55

## 2) INFORMATION FOR SEQ ID NO: 735

## (i) SEQUENCE CHARACTERISTICS:

```

60 (A) LENGTH: 825 bases

```

(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Serratia grimesii*  
 (B) STRAIN: ATCC 14460

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 735

	GGCGCTATCC	TGGTTGTTGC	TGCGACTGAT	GGCCCAATGC	CACAGACCCG	50
	TGAGCACATC	CTGCTGGGTC	GTCAGGTTGG	CGTTCCTTTC	ATCATCGTAT	100
15	TCATGAACAA	ATGCGACATG	GTTGATGATG	AAGAGCTGCT	GGAAGTGGTA	150
	GAAATGGAAG	TTCGTGAACT	TCTGTCTGCT	TATGACTTCC	CTGGTGATGA	200
	CCTGCCAGTT	GTTCGTGGTT	CAGCGCTGAA	AGCACTGGAA	GGCGAAGCTG	250
	AGTGGGAAGC	TAAAATCATC	GAAGTGGCTG	GCTACCTGGA	TTCTTACATC	300
	CCAGAACCAG	AGCGTGCTAT	CGACAAGCCG	TTCTTGCTGC	CAATCGAAGA	350
20	CGTATTCTCC	ATCTCCGGYC	GTGGTACYGT	AGTTACCGGT	CGTGTAGAGC	400
	GCGGTATCGT	TAAAGTTGGC	GAAGAAGTTG	AAATCGTTGG	TATCAAAGAC	450
	ACCGTTAAGT	CTACCTGTAC	TGGCGTTGAA	ATGTTCCGCA	AACTGCTGGA	500
	CGAAGGCCGT	GCTGGTGAGA	ACGTAGGTGT	TCTGCTGCGT	GGTATCAAGC	550
	GTGAAGACAT	CGAACGTGGT	CAGGTACTGG	CTAAACCAGG	TTCAATCAAG	600
25	CCACACACCA	AATTCGACTC	AGAAGTTTAC	ATCCTGAGCA	AAGAAGAAGG	650
	TGGTCGTCAC	ACTCCATTCT	TCAAAGGCTA	CCGTCCACAG	TTCTACTTCC	700
	GTACAACCTGA	CGTGACCGGT	ACCATCGAAC	TGCCAGAAGG	CGTAGAGATG	750
	GTAATGCCAG	GCGATAACGT	GAACATGGTT	GTAACCCTGA	TTCACCCAAT	800
30	CGCGATGGAC	GACGGTCTGC	GTTTC			825

## 2) INFORMATION FOR SEQ ID NO: 736

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 798 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium difficile*  
 (B) STRAIN: ATCC 9689

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 736

	TATTAGTTTG	TTCAGCAACA	GATGGACCAA	TGCCACAAAC	AAGAGAGCAT	50
50	ATACTATTAT	CAAGACAAGT	TGGAGTACCA	TATATAGTAG	TATTCTTAAA	100
	CAAATGTGAC	ATGGTAGATG	ATGAAGAGTT	ATTAGAGTTA	GTAGAGATGG	150
	AAGTAAGAGA	TTTATTAACA	GAATATGATT	TCCCAGGAGA	TGACACTCCA	200
	ATAGTAAGAG	GTTTCAGCATT	AATGGCATT	GAAGATCCAA	AGAGYGAGTG	250
	GGGAGATAAG	ATAGTAGAAT	TATTCGAGCA	AATAGATGAG	TATATACCAG	300
55	CTCCAGAGAG	AGATACAGAT	AAACCATCT	TAATGCCAGT	AGAGGACGTA	350
	TTCTCAATCA	CAGGAAGAGG	AACAGTTGCA	ACAGGAAGAG	TGGAAAGAGG	400
	AGTACTAAAA	GTACAAGACG	AAGTAGAKTT	AGTAGGATTA	ACAGAAGCAC	450
	CAAGAAAAGT	AGTAGTAACA	GGAGTAGAGA	TGTTTCAGAAA	ATTATTAGAC	500
	CAAGCACAAAG	CAGGGGATAA	TATAGGAGCA	TTATTAAGAG	GAGTACAAAG	550
60	AAACGAGATA	GAAAGAGGAC	AAGTACTAGC	AAAGACTGGA	TCAGTAAAGG	600

CACACACAAA	GTTTACAGCA	GAAGTATATG	TACTTAAAAA	AGAAGARGGT	650
GGAAGACATA	CACCATTCCT	TGATGGATAT	AGACCACAAT	TCTATTTTCAG	700
AACAACAGAC	GTAACAGGAG	CTTGTAAGTT	ACCAGAAGGA	ATAGAGATGG	750
TAATGCCTGG	AGATAACGTA	ACAATGGAAG	TAGACTTAAT	AAACTCAA	798

5

## 2) INFORMATION FOR SEQ ID NO: 737

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 411 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: *Burkholderia pseudomallei*  
 (B) STRAIN: 1026B

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 737

25	GTCAACATGA	TGGAGCTCAT	CAACAACATC	GCGAAGGAGC	ACGGCGGTTA	50
	CTCCGTGTTT	GCGGGCGTGG	GCGAGCGTAC	CCGTGAAGGG	AACGACTTCT	100
	ACCACGAAAT	GAAGGACTCG	AACGTTCTCG	ACAAGGTCGC	GCTGGTGTAC	150
	GGCCAGATGA	ACGAGCCGCC	GGGCAACCGT	CTGCGCGTGG	CGCTGACGGG	200
	CCTCACGATG	GCCGAGCACT	TCCGTGACGA	AGGCCTCGAC	GTGCTGTTCT	250
	TCGTCGACAA	CATCTACCGT	TTCACGCTGG	CCGGTACCGA	AGTGTCGGCG	300
30	CTGCTCGGCC	GTATGCCGTC	GGCAGTGGGC	TATCAGCCGA	CGCTGGCTGA	350
	AGAAATGGGC	AAGCTGCAAG	AGCGCATCAC	GTGACGAAG	AAGGGCTCGA	400
	TCACGTCGGT	T				411

35

## 2) INFORMATION FOR SEQ ID NO: 738

40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 394 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

45

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium bifermentans*  
 (B) STRAIN: ATCC 638

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 738

55	TACAAGAGCT	TATTAACAAT	ATAGCTACTC	AACACGGTGG	TATATCAGTA	50
	TTTCGAGGTG	TTGGAGAGAG	AACAAGAGAA	GGTAACGACT	TATTCCATGA	100
	GATGAGCGAT	ACAGGAGTTA	TAAATAAAAC	AGCTCTAGTA	TTCGGACAAA	150
	TGAATGAGCC	ACCTGGAGCA	AGAATGAGAG	TTGCTTTAAC	TGGTCTTACA	200
	ATGGCTGAAT	ACTTCAGAGA	TCAACAAGGG	CAAGACGTTT	TATTATTCGT	250
	AGATAATATA	TTCCGTTTCA	CTCAAGCAGG	ATCTGAGGTT	TCTGCACTTC	300
	TTGGACGTAC	TCCATCAGCA	GTTGGATACC	AACCAACATT	AGCAACAGAG	350
60	ATGGGTAGAT	TACAAGAGAG	AATAACATCT	ACAAATAAAG	GGTC	394

## 2) INFORMATION FOR SEQ ID NO: 739

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 394 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Clostridium beijerincki* (deposited as  
 15 *Clostridium butyricum*)  
 (B) STRAIN: ATCC 8260

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 739

20	TTAATAAACA ACATAGCTAA ACAACATGGT GGTTTATCAG TATTTACTGG	50
	AGTTGGTGAA AGATCAAGAG AAGGTAATGA CTTATATCAT GAAATGAGAG	100
	AGTCAGGAGT TATTGATAAG ACAGCATTAG TATTTGGACA AATGAATGAG	150
	CCACCGGGTG CCAGAATGAG AGTTGCATTA ACAGGTCTTA CTATGGCAGA	200
	GTATTTTAGA GATAAAGGTC AAGATGTGTT ACTATTCATA GATAACATAT	250
25	TCAGATATAC TCAAGCAGGT TCAGAGGTTT CAGCATTACT TGGAAGAACA	300
	CCTTCAGCGG TTGGATATCA GCCAACACTT GCAACTGAAA TGGGTGCACT	350
	TCAGGAAAGA ATTACATCAA CAGTTAATGG TTCTATTACG TCAG	394

30

## 2) INFORMATION FOR SEQ ID NO: -740 -----

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 393 bases  
 35 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

40 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Clostridium difficile*  
 (B) STRAIN: ATCC 9689

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 740

50	TTATAAACAA TATTGCTAAG CAACATGGTG GTATTTCTGT ATTTTCAGGA	50
	GTAGGAGAAA GAACAAGAGA AGGTAACGAC CTTTATGGCG AAATGAGTGA	100
	GTCTGGAGTT ATAAATAAAA CAGCTCTAGT ATTTGGTCAA ATGAATGAAC	150
	CACCTGGAGC GAGAATGAGA GTTGCTTTAA CTGGACTTAC AATGGCAGAA	200
	CATTTTAGAG ATGAGCAAGG ACAAGACGTT TTA CTTTTCG TTGATAATAT	250
	ATTCCGTTTC ACACAAGCTG GTTCAGAAGT TTCAGCACTT CTAGGACGTA	300
	TGCCATCAGC TGTTGGTTAT CAGCCAACAT TAGCTACTGA AATGGGTGCA	350
55	CTTCAAGAGA GAATAACATC AACTAAGAAA GGTTCAATAA CAT	393

55

## 2) INFORMATION FOR SEQ ID NO: 741

60 (i) SEQUENCE CHARACTERISTICS:

625



- (A) LENGTH: 398 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium ramosum*
- (B) STRAIN: ATCC 25582

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 741

15	TTGATTCAAG AATTCATTAA TAACATTGCT ACAGAACATG GTGGTTTATC	50
	AGTTTTTGCT GGAGTTGGTG AACGTAGCCG TGAAGGTAAT GATTTATATT	100
	ATGAAATGAA GGAAAGTGGT GTTTTATCTA AAACAACACT AGTATTTGGA	150
	CAGATGAATG AACCCCCAGG AGCTCGTTTA AGAGTTGCTT TAACGGGTCT	200
	TACTATGGCA GAAGAATTCC GTGATGAACA AGGTCAGGAT GTCTTATTAT	250
	TCATCGATAA TATTTTCCGT TTTACTCAAG CTGGATCTGA AGTATCTGCC	300
20	TTACTTGGAC GGGTACCATC ACAAGCTGGG TATCAGCCAA CTTTAGCAAC	350
	CGAAATGGGT GCTTTACAAG AACGGATTAC ATCAACTAAA AAAGGATC	398

25 2) INFORMATION FOR SEQ ID NO: 742

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium septicum*
- (B) STRAIN: ATCC 12964

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 742

40	TAGCTAAGGA ACACGGTGGA CTTTCAGTAT TCACAGGTGT TGGAGAAAGA	50
	TCAAGAGAAG GTAATGATTT ATATTACGAA ATGAAAGAAT CAGGAGTTAT	100
	AGACAAGACA GCTCTAGTGT TTGGACAAAT GAATGAATCT CCAGGAGCTA	150
	GAATGAGAGT ATCTTTAACA GGATTAACATA TGGCTGAATA TTTCAGAGAT	200
45	CAAGGTCAAG ATGTGCTTTT ATTCATAGAT AACATATTTA GATTTACTCA	250
	AGCTGGATCA GAAGTATCGG CTTTACTTGG AAGAATACCA TCAGCAGTTG	300
	GTTATCAACC AACACTAGCA ACTGAAATGG GTGCACTTCA AGAAAGAATT	350
	ACTTCAACTA AAAATGGATC AATAACTTCA	380

50

2) INFORMATION FOR SEQ ID NO: 743

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 389 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

55

(ii) MOLECULE TYPE: Genomic DNA

60

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium tertium*

(B) STRAIN: ATCC 14573

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 743

5  
10  
15

TTAATAAATA	ATATAGCAAA	AGAGCATGGT	GGTCTTTCTG	TATTTACAGG	50
AGTTGGAGAA	AGGTCAAGAG	AAGGTAACGA	CTTATATTAT	GAAATGAAAG	100
AGTCAGGGGT	TATAGATAAG	ACAGCTTTAG	TATTTGGACA	AATGAATGAA	150
TCACCAGGAG	CAAGAATGAG	AGTTTCATTA	ACTGGATTAA	CTATGGCTGA	200
ATATTTTAGA	GATCAAGGTC	AAGACGTTCT	TTTATTTATA	GATAATATAT	250
TTAGATTTAC	TCAAGCGGGA	TCAGAAGTTT	CTGCGTTATT	AGGAAGAATT	300
CCTTCAGCAG	TTGGATATCA	ACCAACTCTT	GCAACTGAAA	TGGGAGCACT	350
TCAAGAAAGA	ATAACATCAA	CAAAGAATGG	ATCAATCAC		389

## 2) INFORMATION FOR SEQ ID NO: 744

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 843 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Comamonas acidovorans*

(B) STRAIN: ATCC 15668

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 744

35  
40  
45  
50

TTCCCCCGCA	CGCATGCCCA	AGGTGTTCGA	TGCCCTGAAG	CTCGACGGCT	50
CGGCCCTGAC	GCTGGAAGTG	CAGCAACTGC	TGGGTGACGG	CGTTGTGCGT	100
ACCATCGCCC	TGGGTTTCGTC	CGACGGTCTG	CGTCGCGGCC	TGATGGTGTC	150
CAACACCGGC	AACCCCATCA	CCGTGCCCGT	GGGCAAGGCG	ACGCTGGGTC	200
GCATCATGGA	CGTGCTGGGC	AATCCCATCG	ACGAACGTGG	TCCCGTGGAT	250
CAGGCGCTGA	CGGCTCCCAT	CCACCGCAAG	GCACCGGCTT	ATGACGAGCT	300
GTCGCCTTCG	CAGGAACTGC	TGGAAACCGG	CATCAAGGTG	ATCGACCTGA	350
TCTCGCCCTT	CGCCAAGGGC	GGCAAGGTGG	GTCTGTTCGG	TGGCGCCGGT	400
GTGGGCAAGA	CCGTGAACAT	GATGGAACTC	ATCAACAACA	TCGCCAAGGG	450
CCACGGTGGT	CTGTGCGGTG	TCGCCGGTGT	GGGTGAACGT	ACCCGCGAAG	500
GCAATGACTT	CTATCACGAA	ATGTCGGACG	CCGGCGTGGT	CAACCAGGAG	550
TCGCTGAACG	ACTCCAAGGT	GGCCATGGTC	TACGGCCAGA	TGAACGAACC	600
CCCGGGCAAC	CGTCTGCGCG	TGGCGCTGAC	CGGCCTGACC	ATGGCCGAAG	650
CCTTCCGTGA	CGAAGGCAAG	GACGTGCTGT	TCTTCGTGGA	CAACATCTAC	700
CGCTACACGC	TGGCCGGTAC	CGAAGTGTCC	GCTCTGCTGG	GTCGCATGCC	750
TTCCGCCGTG	GGCTACCAGC	CCACGCTGGC	CGAGGAAATG	GGCCGCCTGC	800
AAGAGCGCAT	CACCTCGACC	AAGGTCGGTT	CGATCACTTC	CAC	843

## 2) INFORMATION FOR SEQ ID NO: 745

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 819 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Klebsiella pneumoniae* subsp.  
*rhinoscleromatis*

(B) STRAIN: ATCC 13884

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 745

	GCCGTACCAC	GCGTGACGA	AGCCCTTGAG	GTACAGAATG	GTAATGAAGT	50
	TCTGGTGCTG	GAAGTTCAGC	AGCAGCTGGG	CGGCGGTATC	GTACGTACCA	100
	TCGCCATGGG	TTCTTCTGAT	GGTCTGCGCC	GCGGTCTGGA	TGTAAAAGAC	150
15	CTCGAGCACC	CGATCGAAGT	CCCGGTAGGT	AAAGCAACGC	TGGGTCGTAT	200
	CATGAACGTA	CTGGGTCAAC	CGGTGACAT	GAAAGGCGAC	ATCGGCGAAG	250
	AAGAGCGTTG	GGCTATCCAC	CGCGCGGCAC	CGTCCTATGA	AGAGCTGTCC	300
	AGCTCTCAGG	AACTGCTGGA	AACCGGCATC	AAAGTTATCG	ACCTGATGTG	350
	TCCGTTTCGCC	AAGGGCGGTA	AAGTTGGTCT	GTTCCGGCGGT	GCGGGTGTAG	400
20	GTAAAACTGT	AAACATGATG	GAGCTGATCC	GTAACATCGC	GATCGAGCAC	450
	TCCGGTTACT	CTGTGTTTGC	GGGCGTAGGT	GAGCGTACTC	GTGAGGGTAA	500
	CGACTTCTAC	CACGAAAATGA	CCGACTCCAA	CGTTATCGAT	AAAGTATCCC	550
	TGGTGTACGG	CCAGATGAAC	GAGCCGCCGG	GAAACCGTCT	GCGCGTTGCG	600
	CTGACCGGCC	TGACCATGGC	TGAGAAATTG	CGTGACGAAG	GTCGTGACGT	650
25	ACTGCTGTTT	GTCGATAACA	TCTATCGTTA	CACCCTGGCC	GGTACTGAAG	700
	TATCCGCGCT	GCTGGGTCGT	ATGCCTTCAG	CGGTAGGTTA	TCAGCCGACC	750
	CTGGCGGAAG	AGATGGGCGT	TCTGCAGGAA	CGTATCACCT	CCACCAAAAC	800
	CGGTTCTATC	ACCTCCGTA				819

30

2) INFORMATION FOR SEQ ID NO: 746

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 824 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

40 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria canis*

(B) STRAIN: ATCC 14687

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 746

	GCGATTCTAT	TCCGCGCGTA	TATGATGCTC	TTAAACTAGT	GGATAGAGAA	50
	CTGACGCTTG	AAGTACAACA	ACAGTTGGGT	GATGGTGTCG	TTCGTACTAT	100
50	TGCGATGGGT	AGTTCCGACG	GCCTCAAACG	AGGTTTGCGG	GTAGTTAACA	150
	CCGGTGCTCC	AATTACAGTG	CCTGTGGGCA	AAGCAACATT	AGGCCGTATT	200
	ATGGACGTAT	TAGGTAATCC	GGTTGATGAA	GCTGGTCCGA	TTGGCTCCGA	250
	GCAAACCCGA	GCAATCCACC	AACCTGCTCC	TAAGTTCGAC	GAGCTTTCTA	300
	GCGCCACAGA	GATTTTGGAA	ACAGGTATTA	AAGTAATTGA	TTTGCTTTGC	350
55	CCGTTTGCCA	AAGGCGGTAA	AGTAGGTTTG	TTTGGTGGTG	CGGGAGTGGG	400
	CAAAACCGTA	AATATGATGG	AGTTGATTAA	CAACATCGCG	AAAGCACACA	450
	GCGGTTTGTC	TGTATTTGCC	GGTGTGGGTG	AACGGACGCG	TGAAGGTAAT	500
	GACTTTTATC	ATGAGATGAA	AGATTCCAAT	GTATTAGATA	AAGTAGCCAT	550
	GGTTTACGGT	CAGATGAACG	AGCCTCCCGG	TAACCGTTTG	CGCGTTGCGC	600
60	TAAGTGGCTT	GTCTATGGCC	GAATTCTTCC	GTGACGAGAA	AGATGAAAAC	650

GGTAAAGGCC	GTGATGTATT	GTTCTTTGTA	GACAATATTT	ACCGCTATAC	700
CCTAGCCGGT	ACAGAAGTAT	CTGCATTGCT	TGGCCGTATG	CCTTCGGCAG	750
TAGGTTATCA	GCCGACGTTG	GCAGAGGAAA	TGGGCCGCTT	GCAAGAGCGT	800
ATTACYTCAM	CCCAAACAGG	CTCT			824

5

## 2) INFORMATION FOR SEQ ID NO: 747

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 831 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 15 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Neisseria cinerea*  
 (B) STRAIN: ATCC 14685
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 747

CGCGACGCTA	TCCCGCATGT	TTACGATGCC	CTGAAATTGG	ACGAGAACGG	50
TCTGACTCTG	GAGGTTCAAC	AGCTTCTGGG	CGACGGCGTT	GTCCGTACTA	100
TTGCAATGGG	TAGTTCAGAC	GGCCTTAAAC	GCGGTATGTC	TGTAAGCAAT	150
ACTGGTGCGC	CAATCACTGT	GCCGGTAGGT	AAAGGTACAT	TGGGTCGTAT	200
TGTCGACGTA	TTGGGTACGC	CTGTTGATGA	AGCAGGTCCG	ATCGATACCG	250
ACAAAAGCCG	TGCCATTAC	CAAACCTGCTC	CGAAATTCGA	CGAGTTGTCT	300
30 TCAGCTACCG	AATTGTTGGA	AACCGGTATT	AAAGTGATCG	ACTTGCTGTG	350
TCCGTTTGCT	AAAGGCGGTA	AAGTAGGTCT	GTTGCGGTGGT	GCCGGTGTGG	400
GCAAAACCGT	GAACATGATG	GAATTGATCA	ACAACATCGC	CAAAGCGCAC	450
AGCGGTCTGT	CCGTGTTTCG	AGGTGTGGGT	GAGCGTACCC	GTGAAGGTAA	500
CGACTTCTAC	CACGAGATGA	AAGATTCCAA	CGTATTGGAT	AAAGTAGCCA	550
35 TGGTGATATG	CCAAATGAAC	GAACCTCCGG	GCAACCGTCT	GCGCGTTGCT	600
TTGACCGGTT	TGACTATGGC	CGAATACTTC	CGTGACGAAA	AAGACGAAAA	650
CGGTAAAGGC	CGCGACGTAT	TGTTCTTCGT	TGACAACATC	TACCGTTACA	700
CTTTGGCCGG	TACTGAAGTA	TCTGCACTGT	TGGGCCGTAT	GCCTTCTGCA	750
GTGGGTTACC	AACCGACATT	GGCTGAAGAA	ATGGGTCGTT	TGCAAGAGCG	800
40 TATTACCTCT	ACCCAAACCG	GTTCCATTAC	T		831

## 2) INFORMATION FOR SEQ ID NO: 748

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 862 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 50 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Neisseria cuniculi*  
 (B) STRAIN: ATCC 14688
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 748

60 CCGTGGCCAA	GTACCACAAA	TTTATGACGC	ACTGAGTGTT	GATGGCACCG	50
---------------	------------	------------	------------	------------	----

```

AAACAACCTT GGAAGTTCAA CAGCAGTTGG GTGATGGCGT GGTGCGTACC 100
ATTGCGATGG GTTCAACCGA AGGCTTGAAG CGTGGTTTGA CTGTATCTAA 150
CTCTGGTGCA CCGATTTCTG TGCCAGTGGG TCAAGCGACT TTGGGTCGTA 200
TTATGGATGT GTTGGGTCGT CCAATCGACG AGGCAGGTCC TGTAATGCT 250
5 CAAGAAAAAT GGTC AATTCA CCGTGAAGCA CCAAGCTATG ATGAGCAATC 300
AAACTCAACT GAGCTGCTAG AAACAGGCAT CAAAGTAATT GATTTGCTTT 350
GCCCATTTGC TAAAGGTGGT AAAGTTGGCT TGTTCGGTGG TGCAGGTGTG 400
GGTAAAACCG TGAATATGAT GGAGCTGATT AATAATATCG CTCTGAAGCA 450
TTCAGGTCTT TCTGTTTTTG CAGGTGTTGG TGAGCGTACT CGTGAGGGTA 500
10 ACGATTTTTA TCACGAAATG CAAGAAGCAG GCGTTGTTAA TACCGAAGAC 550
TTCACCAAGT CAAAAGTAGC GATGGTTTAT GGTCAGATGA ATGAGCCACC 600
AGGAAACCGT TTGCGTGTTG CATTGACAGG CTTGACGATG GCAGAATATT 650
TCCGTGACGA AAAAGATGAA GCAACAGGCA AAGGGCGTGA TGTTCTATTG 700
TTCGTTGATA ACATCTATCG TTACACACTG GCTGGTACGG AAGTGTCAGC 750
15 ACTTCTAGGT CGTATGCCAT CAGCAGTAGG TTATCAACCG ACTCTGGCTG 800
AAGAAATGGG TGC GTTCAA GAGCGTATTA CCTCAACGCA ATCGGGTTCC 850
ATCACTTCGG GG 862

```

```

20 2) INFORMATION FOR SEQ ID NO: 749

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 844 bases
        (B) TYPE: Nucleic acid
25    (C) STRANDEDNESS: Double
        (D) TOPOLOGY: Linear

    (ii) MOLECULE TYPE: Genomic DNA
30    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Neisseria elongata subsp. elongata
        (B) STRAIN: ATCC 25295

```

```

35    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 749

GGAAGTCCCA CGTGACGCTA TCCCGCATGT TTTTGATGCA TTAAAATTAG 50
TTGAAAATGA CCTAACCTTA GAAGTTCAAC AACTTTTGGG GGATGGTGTA 100
GTGCGTACCA TTGCGATGGG TAGTTCAGAT GGATTAAAGC GTGGTATGGC 150
40 TGTGAATAAT ACCGGAGCTC CGATTACTGT TCCTGTTGGC CGTGAAACTT 200
TGGGTCGTAT CATGGATGTA TTGGGTAATC CGGTTGATGA GGCAGGTCCG 250
GTAAATGCAT CCAATACACG TGCGATCCAT CAAGAGGCTC CTAAGTTTGA 300
TGAGCTTTCT TCAACAACGG AATTATTAGA AACTGGCATT AAGGTTATCG 350
ACTTGTTATG TCCGTTTGCC AAAGGTGGTA AAGTAGGTCT GTTTGGTGGT 400
45 GCGGGTGTAG GTAAAACCGT AAATATGATG GAGTTAATTA ACAACATTGC 450
CAAGGCACAT AGCGGTTTGT CTGTGTTTGC AGGCGTGGGT GAACGTACTC 500
GTGAAGGTAA TGACTTCTAT CACGAGATGA AAGATTCCAA CGTATTGGAC 550
AAAGTGGCAA TGGTTTACGG TCAGATGAAC GAACCTCCAG GCAACCGTCT 600
GCGCGTTGCT TTGACCGGTT TGACTATGGC CGAATACTTC CGTGACGAAA 650
50 AAGACGAAAA CGGTAAAGGT CGCGACGTAT TGTCTTTCGT GGACAACATT 700
TACCGTTACA CTTTGCCCGG TACGGAAGTA TCCGCATTGC TGGGTCGTAT 750
GCCTTCAGCA GTAGGTTACC AACCGACATT GGCTGAAGAA ATGGGTCGTT 800
TGCAAGAGCG TATTACCTCT ACCCAGACAG GCTCTATTAC TTCC 844

```

```

55 2) INFORMATION FOR SEQ ID NO: 750

```

```

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 834 bases
60

```

(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria flavescens*  
 (B) STRAIN: ATCC 13120

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 750

	CGCGACGCTA	TTCCGCATGT	TTACGATGCC	CTGAAATTGG	ACGAGAACGG	50
	TCTGACTCTG	GAAGTTCAAC	AGCTTCTGGG	TGACGGCGTT	GTCCGTACTA	100
15	TTGCAATGGG	TAGTTCAGAC	GGCCTGAAAC	GCGGCATGTC	TGTAAGCAAT	150
	ACCGGTGCGC	CAATCACTGT	GCCGGTAGGT	AAAGGTACAT	TGGGCCGTAT	200
	TGTTCGACGT	TTGGGTACGC	CTGTTGATGA	AGCAGGTCCG	ATCGATACCG	250
	ACAAGAGCCG	TGCCATTAC	CAAACGTGCTC	CGAAATTTCGA	TGAGTTGTCT	300
	TCAACTACCG	AATTGTTGGA	AACCGGTATT	AAAGTGATCG	ACTTGCTGTG	350
20	TCCGTTTGCT	AAGGGCGGTA	AAGTAGGTCT	GTTCGGTGGT	GCCGGTGTGG	400
	GCAAAACCGT	GAACATGATG	GAATTGATCA	ACAACATCGC	CAAAGCGCAC	450
	AGCGGCCTGT	CCGTGTTTCG	AGGTGTGGGT	GAACGTACCC	GTGAAGGTAA	500
	CGACTTCTAC	CACGAGATGA	AAGATTCCAA	CGTATTGGAT	AAAGTAGCCA	550
	TGGTGTATGG	TCAAATGAAC	GAACCTCCAG	GCAACCGTCT	GCGCGTTGCT	600
25	TTGACCGGTT	TGACTATGGC	CGAATACTTC	CGTGACGAAA	AAGATGAAAA	650
	CGGTAAAGGT	CGCGACGTAT	TGTTCTTCGT	TGACAACATC	TACCGTTACA	700
	CTCTGGCCGG	TACCGAAGTA	TCCGCACTGT	TGGGTCGTAT	GCCTTCTGCA	750
	GTGGGTTACC	AACCGACATT	GGCTGAAGAA	ATGGGTCGTT	TGCAGGAGCG	800
30	TATTACTTCT	ACGCAAACCG	GTTCCATTAC	TTCC		834

2) INFORMATION FOR SEQ ID NO: 751

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 834 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria gonorrhoeae*  
 (B) STRAIN: ATCC 31426

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 751

	CGCGACATGA	TTCCGCGCGT	TTACGACGCT	TTGAAATTAG	ACGAAAACGG	50
50	TCTGACTTTG	GAAGTCCAAC	AGCTTTTGGG	TGATGGCGTA	GTCCGTACCA	100
	TCGCTATGGG	CAGCTCGGAC	GGTTTGAAAC	GCGGCATGAC	TGTGAGCAAT	150
	ACTGGTTCGC	CCATTACTGT	GCCGGTAGGT	AAAGGTACGT	TGGGACGCAT	200
	TGTTCGATGA	TTGGGAACTC	CTGTTGACGA	GGCAGGTCCA	ATTGATACCG	250
	ACAAGAGTCG	TGCCATCCAC	CAAGCCGCTC	CTAAGTTTGA	CGAACTGTCT	300
55	TCCACAACCG	AATTGCTCGA	AACGGGCATT	AAAGTGATTG	ACTTGCTGTG	350
	TCCGTTTGCC	AAAGGCGGTA	AAGTAGGTCT	GTTCGGCGGT	GCCGGTGTGG	400
	GTAAAACCGT	GAACATGATG	GAATTGATCA	ACAACATCGC	CAAAGCGCAC	450
	AGCGGCTTGT	CCGTGTTCTC	AGGCGTGGGT	GAGCGTACCC	GCGAAGGTAA	500
	CGACTTCTAC	CACGAGATGA	AAGATTCCAA	CGTATTGGAT	AAAGTAGCCA	550
60	TGGTGTATGG	CCAAATGAAC	GAACCTCCAG	GCAACCGTCT	GCGCGTTGCT	600

	TTGACCGGTT	TGACTATGGC	TGAATACTTC	CGTGACGAAA	AAGACGAAAA	650
	CGGTAAAGGT	CGTGACGTAT	TGTTCTTCGT	TGACAACATC	TACCGTTACA	700
	CTCTGGCCCG	TACCGAAGTA	TCCGCACTGT	TGGGCCGTAT	GCCTTCTGCA	750
	GTGGGTTACC	AACCGACATT	GGCTGAAGAA	ATGGGTCGTT	TGCAAGAGCG	800
5	TATTACCTCT	ACCCAAACCG	GTTCCATTAC	TTCC		834

## 2) INFORMATION FOR SEQ ID NO: 752

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 825 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria gonorrhoeae*
- (B) STRAIN: ATCC 27628

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 752

25	ACATGATTCC	GCGCGTTTAC	GACGCTTTGA	AATTAGACGA	AAACGGTCTG	50
	ACTTTGGAAG	TCCAACAGCT	TTTGGGTGAT	GGCGTAGTCC	GTACCATCGC	100
	TATGGGCAGC	TCGGACGGTT	TGAAACGCGG	CATGACTGTG	AGCAATACTG	150
	GTTGCCCCAT	TACTGTGCCG	GTAGGTAAAG	GTACGTTGGG	ACGCATTGTC	200
	GATGTATTGG	GAACGCCTGT	TGACGAGGCA	GGTCCAATTG	ATACCGACAA	250
30	GAGTCGTGCC	ATCCACCAAG	CCGCTCCTAA	GTTTGACGAA	CTGTCTTCCA	300
	CAACCGAATT	GCTCGAAACG	GGCATTAAAG	TGATTGACTT	GCTGTGTCCG	350
	TTTGCCAAAG	GCGGTAAAGT	AGGTCTGTTC	GGCGGTGCCG	GTGTGGGTAA	400
	AACCGTGAAC	ATGATGGAAT	TGATCAACAA	CATCGCCAAA	GCGCACAGCG	450
	GCTTGTCGGT	GTTCTCAGGC	GTAGGTGAGC	GTACCCGCGA	AGGTAACGAC	500
35	TTCTACCACG	AGATGAAAGA	TTCCAACGTA	TTGGATAAAG	TAGCCATGGT	550
	GTATGGCCAA	ATGAACGAAC	CTCCAGGCAA	CCGTCTGCGC	GTTGCTTTGA	600
	CCGGTTTGAC	TATGGCTGAA	TACTTCCGTG	ACGAAAAGA	CGAAAACGGT	650
	AAAGGTCGTG	ACGTATTGTT	CTTCGTTGAC	AACATCTACC	GTTACACTCT	700
	GGCCGGTACC	GAAGTATCCG	CACTGTTGGG	CCGTATGCCT	TCTGCAGTGG	750
40	GTTACCAACC	GACATTGGCT	GAAGAAATGG	GTCGTTTGCA	AGAGCGTATT	800
	ACCTCTACCC	AAACCGGTTC	CATTA			825

## 2) INFORMATION FOR SEQ ID NO: 753

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 831 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria lactamica*
- (B) STRAIN: ATCC 23970

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 753

```

5  GATGCGATTTC CGCATGTTTA CGATGCCCTG AAATTGGACG AGAACGGTCT 50
   GACCCTGGAA GTCCAACAGC TTTTGGGTGA CGGCGTTGTC CGTACTATTG 100
   CAATGGGTAG TTCAGACGGC CTGAAACGCG GCATGTCTGT CAGCAATACC 150
   GGTGCGCCAA TCACTGTGCC GGTAGGTAAA GGTACGTTGG GCCGTATTGT 200
   CGACGTATTG GGTACACCTG TTGACGAAGC AGGTCCGATC GATACCGACA 250
   AGAGCCGCGC CATCCACCAA ACCGCCCGCA AATTCGACGA GTTGTCTTCA 300
   ACTACCGAAT TGTGGAAC CGGCATTAAA GTGATCGATT TGCTGTGTCC 350
   GTTTGCTAAG GGCGGTAAAG TAGGTCTGTT CGGTGGTGCC GGTGTGGGCA 400
   AAACCGTGAA CATGATGGAA TTGATCAACA ACATCGCCAA AGCGCACAGC 450
10 GGTCTGTCCG TGTTTGCAGG CGTGGGCGAG CGTACCCGCG AAGGTAACGA 500
   CTTCTACCAC GAGATGAAAG ATTCCAACGT ATTGGATAAA GTAGCCATGG 550
   TGTATGGTCA GATGAACGAA CCTCCGGGCA ACCGTCTGCG CGTTGCTTTG 600
   ACCGGTTTGA CGATGGCCGA ATACTTCCGC GACGAAAAG ACGAAAACGG 650
   CAAAGGCCGC GACGTATTGT TCTTCGTGGA CAACATCTAC CGTTACACCC 700
15 TGGCCGGTAC CGAAGTATCC GCACTGTTGG GCCGTATGCC TTCCGCAGTG 750
   GGTTACCAAC CGACATTGGC TGAAGAAATG GGTTCGTTTG C AAGAGCGTAT 800
   TACCTCTACC CAAACCGGTT CCATTACTTC C 831

```

20

## 2) INFORMATION FOR SEQ ID NO: 754

## (i) SEQUENCE CHARACTERISTICS:

```

25 (A) LENGTH: 836 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
   (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

30

## (vi) ORIGINAL SOURCE:---

```

   (A) ORGANISM: Neisseria meningitidis
   (B) STRAIN: 2241C

```

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 754

```

   CCACGCGACA TGATTCCGCG CGTTTACGAC GCTTTGAAAT TAGACGAAAA 50
   CGGTCTGACT TTGGAAGTCC AACAGCTTTT GGGCGACGGC GTAGTCCGTA 100
   CCATTGCGAT GGGCAGCTCG GACGGTTTGA AACCGGCGCAT GACTGTGAGC 150
40 AATACCGGTG CGCCCATTAC TGTGCCGGTA GGTAAAGGTA CGTTGGGACG 200
   CATTGTGCGAT GTATTGGGAA CGCCTGTTGA CGAGGCAGGT CCAATCGATA 250
   CCGACAAGAG CCGTGCCATC CACCAAGCCG CTCTAAGTT TGACGAAC TG 300
   TCTTCCACAA CCGAATTGCT CGAAACGGGC ATTAAAGTGA TTGACTTGCT 350
   GTGTCCGTTT GCCAAAGGCG GTAAAGTAGG TCTGTTCGGC GGTGCCGGTG 400
45 TGGGTAAAAC CGTGAACATG ATGGAATTGA TCAACAACAT CGCCAAAGCG 450
   CACAGCGGCT TGTCCGTGTT CGCAGGCGTG GGTGAGCGTA CCCGCGAAGG 500
   TAACGACTTC TACCACGAGA TGAAAGATTC CAACGTATTG GATAAAGTGG 550
   CAATGGTTTA CGGTCAGATG AACGAACCTC CGGGCAACCG TTTGCGCGTC 600
   GCATTGACCG GTTTGACCAT GCGGAATAC TTCCGTGACG AAAAAGACGA 650
50 AAACGGCAAA GGCCGCGACG TATTGTTCTT CGTGACAAC ATCTACCGTT 700
   ACACTCTAGC TGGTACCGAA GTATCCGCAT TGTGCGGCCG TATGCCGTCT 750
   GCAGTGGGCT ACCAACCGAC ATTGGCAGAA GAAATGGGTC GTTTGCAGGA 800
   GCGTATTACC TCTACCCAAA CCGGTTCTAT TACTTC 836

```

55

## 2) INFORMATION FOR SEQ ID NO: 755

## (i) SEQUENCE CHARACTERISTICS:

```

60 (A) LENGTH: 837 bases

```



(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria mucosa*  
 (B) STRAIN: ATCC 19696

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 755

	CCGCGTGATG	CCATTCCGCG	TGTTTACGAC	GCCCTGAAAT	TGGATGCAAA	50
	CGGCCTGACT	TTGGAAGTGC	AACAGCTTCT	GGGTGACGGC	GTAGTTCGTA	100
15	CTATTGCAAT	GGGTAGTTCG	GATGGTTTGA	AACGCGGCAT	GA CTGTAAGC	150
	AATACAGGTG	CGCCGATTAC	AGTACCGGTA	GGTAAAGGTA	CTTTGGGACG	200
	TATTGTTCAT	GTATTGGGTA	CGCCTGTTGA	CGAAGCAGGT	CCGATTGATA	250
	CCGACAAACA	CCGTGCCATC	CATCAGACAG	CTCCGAAATT	TGATGAACTG	300
	TCTGCGACTA	CTGAGCTGCT	GGAACCGGC	ATTAAAGTGA	TTGACTTGCT	350
20	GTGTCCGTTT	GCCAAAGGCG	GTAAAGTAGG	TCTGTTCGGT	GGTGCCGGTG	400
	TAGGCAAAAC	CGTCAACATG	ATGGAATTAA	TTAACAACAT	CGCCAAAGCA	450
	CATAGCGGTT	TGTCCGTGTT	TGCAGGTGTG	GGTGAGCGTA	CCCGTGAAGG	500
	TAATGACTTC	TACCACGAGA	TGAAAGATTC	CAACGTATTG	GACAAAGTGG	550
	CGATGGTTTA	CGGTACAGAT	AACGAACCTC	CGGGTAACCG	TCTGCGTGTA	600
25	GCCTTGACCG	GTTTGACGAT	GGCCGAATAC	TTCCGTGATG	AAAAAGACGA	650
	AAGCGGCAAA	GGCCGTGACG	TATTGTTCTT	CGTGGACAAT	ATCTACCGTT	700
	ACACCCTGGC	CGGTACCGAA	GTATCCGCAT	TGTTGGGTTCG	TATGCCTTCA	750
	GCAGTAGGTT	ACCAACCGAC	ATTGGCTGAA	GAAATGGGTC	GTTTGCAAGA	800
30	GCGTATTACC	TCTACCCAAA	CAGGCTCCAT	TACCTCC		837

2) INFORMATION FOR SEQ ID NO: 756

35 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 834 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Neisseria subflava*  
 (B) STRAIN: ATCC 14221

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 756

	CGCGACGCTA	TCCCGCATGT	TTACGATGCC	CTGAAATTGG	ACGAGAACGG	50
50	TCTGACTCTG	GAAGTTCAAC	AGCTTCTGGG	TGACGGCGTT	GTCCGTACTA	100
	TTGCAATGGG	TAGTTCAGAC	GGCCTGAAAC	GCGGCATGTC	TGTAAGCAAT	150
	ACTGGTGCGC	CAATCACTGT	GCCGGTAGGT	AAAGGTACAT	TGGGTCGTAT	200
	TGTCGACGTA	TTGGGTACGC	CTGTTGATGA	AGCAGGTCCG	ATCGATACCG	250
	ACAAGAGCCG	TGCCATTAC	CAAAGTCTC	CGAAATTCGA	CGAGTTGTCT	300
55	TCAACTACCG	AATTGCTGGA	AACCGGTATT	AAAGTGATCG	ACTTGCTGTG	350
	TCCGTTTGCT	AAGGGCGGTA	AAGTAGGTCT	GTTCGGTGGT	GCCGGTGTGG	400
	GCAAAACCGT	GAACATGATG	GAATTGATCA	ACAACATCGC	CAAAGCGCAC	450
	AGCGGTCTGT	CCGTGTTTCG	AGGTGTGGGT	GAACGTACCC	GTGAAGGTAA	500
	CGACTTCTAC	CACGAGATGA	AAGATTCCAA	CGTATTGGAT	AAAGTAGCCA	550
60	TGGTGTATGG	CCAAATGAAC	GAACCTCCGG	GCAACCGTCT	GCGCGTTGCT	600

TTGACCGGTT TGA CTATGGC CGAATACTTC CGTGACGAAA AAGACGAAAA 650  
 CGGTAAAGGT CGCGACGTAT TGTTCTTCGT TGACAACATC TACCGTTACA 700  
 CTCTGGCCGG TACCGAAGTA TCTGCACTGT TGGGCCGTAT GCCTTCTGCA 750  
 GTGGGTTACC AACCGACATT GGCTGAAGAA ATGGGTCGTT TGCAAGAGCG 800  
 5 TATTACCTCT ACCCAAAGT GTTCCATTAC TTCC 834

## 2) INFORMATION FOR SEQ ID NO: 757

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 833 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria weaveri*  
 (B) STRAIN: ATCC 51223

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 757

25 CCCGTGATGC TATTCCCATG TATACGATGC CCTGAAATTG GTAGATAACG 50  
 ATCTGACCCT GGAAGTGCAA CAACTTTTAG GTGATGGTGT GGTTCGTACC 100  
 ATTGCAATGG GTAGTTCAGA CGGCCTAAAA CGTGGTATGG CTGTTAACAA 150  
 TACCGGCGCT CCGATTACTG TTCCGGTGGG GAAAGCCACC TTGGGACGTA 200  
 TTATGGATGT GTTGGGTAAT CCGGTTGATG AAGCAGGTCC TGTTGTATCA 250  
 30 GAAGAACTC GCGCTATTCA TCAAGCTGCC CCTAAATTTG ACGAACTGTC 300  
 TTCAGCAACT GAGTTGTTGG AAACAGGCAT TAAAGTAATT GACTTGCTGT 350  
 GCCCGTTTGC CAAAGGTGGT AAAGTAGGTT TGT TTGGTGG TGCCGGCGTG 400  
 GGTAAAACCG TAAATATGAT GGAGTTGATC AACACATCG CGAAGGCACA 450  
 TAGTGGTTTG TCTGTATTCG CCGGTGTAGG TGAGCGTACC CGTGAAGGTA 500  
 35 ACGACTTCTA CCATGAAATG AAAGACTCTA ACGTATTGGA TAAAGTAGCC 550  
 ATGGTTTATG GCCAGATGAA TGAACCTCCG GGTAAACGTT TGCGCGTTGC 600  
 TTTGACTGGT TTGACTATGG CCGAATATTT CCGTGACGAG AAAGATGAAA 650  
 ACGGCAAAGG TCGTGACGTC TTGTTCTTTG TGGATAATAT CTATCGCTAT 700  
 ACTCTGGCCG GTACTGAAGT GTCTGCACTG TTAGGTCGTA TGCCGTCTGC 750  
 40 AGTAGGTTAT CAGCCTACAT TGGCAGAAGA AATGGGTCGC TTGCAGGAGC 800  
 GTATTACTTC TACTCAAACA GGTTCGATTA CTT 833

## 2) INFORMATION FOR SEQ ID NO: 758

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 833 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria animalis*  
 (B) STRAIN: ATCC 19573

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 758

	CGCGATGCCA	TTCCTCACGT	TTACGATGCC	CTGAAATTGG	ACGACACCGG	50
	TCTGACTTTG	GAAGTACAAC	AACTTCTGGG	CGACGGTGTG	GTACGTACCA	100
	TTGCAATGGG	TAGTTCAGAC	GGCCTGAAAC	GGGGTTTGTC	TGTGAGCAAT	150
	ACCGGTTCTC	CGATTGCCGT	TCCTGTCGGT	AAAGCGACTT	TGGGTCGTAT	200
5	TATGGACGTA	TTGGGCAATC	CGGTTGATGA	AGCCGGTCCG	GTTGCTACCG	250
	AAGAGAAACG	TGCTATTCAC	CAAGCCGCAC	CGAAGTTTGA	TGAATTGTCA	300
	TCAGCTACCG	AGTTGTTGGA	AACCGGTATT	AAAGTAATCG	ACTTGCTGTG	350
	TCCGTTTGCA	AAAGGCGGTA	AAGTAGGTCT	GTTCGGCGGT	GCCGGTGTGG	400
	GCAAAACCGT	AAACATGATG	GAATTGATCA	ACAACATCGC	CAAAGCACAC	450
10	AGCGGTCTGT	CTGTGTTTGC	CGGTGTAGGT	GAACGTACCC	GCGAAGGTAA	500
	CGACTTCTAC	CACGAGATGA	AAGATTCCAA	CGTGTTGGAT	AAAGTAGCCA	550
	TGGTGTACGG	TCAGATGAAT	GAGCCGCCGG	GCAACCGCTT	GCGCGTGGCT	600
	TTGACCGGCC	TGACTATGGC	CGAATACTTC	GTGACGAAAA	AGACGAAAAC	650
	GGCAAAGGTC	GTGACGTATT	GTTCTTCGTG	GACAACATTT	ACCGCTACAC	700
15	ACTGGCCGGT	ACCGAAGTAT	CAGCATTGTT	GGGCCGTATG	CCGTCTGCAG	750
	TAGGTTATCA	GCCGACATTG	GCAGAGGAAA	TGGGTCGCTT	GCAAGAGCGT	800
	ATTACCTCTA	CCCAAACCGG	TTCGATTACC	TCT		833

20

## 2) INFORMATION FOR SEQ ID NO: 759

## (i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 819 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

30

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Proteus penneri*
- (B) STRAIN: ATCC 33519

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 759

	AGCGTCCCTA	AAGTATACGA	CGCTCTTGAG	GTTATGAATG	GTAAAGAAAA	50
	ACTGGTGCTG	GAAGTTCAGC	AACAGTTAGG	CGGTGGTATC	GTTCGTTGTA	100
	TCGCAATGGG	TACATCAGAC	GGTTTAAGCC	GTGGCTTAAA	AGTTGAAGAT	150
40	TTAGGCCACC	CAATTGAAGT	ACCAGTAGGT	AAAGCAACAC	TGGGACGTAT	200
	CATGAACGTT	CTGGGTACAC	CTATTGATAT	GAAAGGTGAT	ATTGCAACTG	250
	AAGAACGTTG	GTCTATTTCAC	CGTGAAGCAC	CAACCTACGA	AGAGTTATCA	300
	AACTCACAAAG	AACTGCTTGA	AACCGGTATC	AAAGTAATGG	ACTTAATCTG	350
	TCCGTTTGCT	AAAGGTGGTA	AAGTAGGTCT	GTTCGGTGGT	GCGGGTGTGG	400
45	GTAAAACAGT	TAACATGATG	GAATTGATCC	GTAATATCGC	GATCGAGCAC	450
	TCAGGTTACT	CTGTATTGTC	TGGTGTTGGT	GAGCGTACTC	GTGAGGGTAA	500
	CGACTTCTAT	CATGAAATGA	CAGATTCTAA	CGTTCTTGAC	AAAGTATCGT	550
	TAGTTTATGG	TCAGATGAAT	GAGCCACCAG	GAAACCGTCT	GCGTGTAGCA	600
	CTGACTGGTC	TGACTATGGC	TGAAAAATTC	CGTGATGAAG	GCCGTGACGT	650
50	ACTGTTATTC	GTGATAACA	TTTATCGTTA	CACCTTAGCC	GGTACAGAAG	700
	TATCAGCACT	GTTAGGTCGT	ATGCCATCAG	CGGTAGGTTA	CCAGCCAACA	750
	TTGGCTGAAG	AGATGGGTGT	TCTGCAAGAA	CGTATCACTT	CAACCAAAC	800
	AGGTTCAATC	ACCTCTGTA				819

55

## 2) INFORMATION FOR SEQ ID NO: 760

## (i) SEQUENCE CHARACTERISTICS:

60

- (A) LENGTH: 819 bases

(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Salmonella choleraesuis* subsp. *choleraesuis*  
 serotype Enteritidis

10 (B) STRAIN: ATCC 13076

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 760

	GCCGTACCAC	GCGTGTACGA	TGCCCTTGAG	GTGCAGAATG	GTAATGAGAA	50
15	GCTGGTGCTG	GAAGTTCAGC	AGCAGCTTGG	CGGCGGTATC	GTGCGTACCA	100
	TCGCGATGGG	GTCTTCTGAC	GGTCTGCGTC	GCGGTCTGGA	TGTAAAAGAT	150
	CTCGAACACC	CGATCGAAGT	CCCGGTAGGT	AAAGCCACGC	TGGGTCGTAT	200
	CATGAACGTC	CTGGGCGAAC	CGGTCGACAT	GAAAGGCGAG	ATCGGCGAAG	250
	AAGAGCGTTG	GGCGATTACAC	CGCGCAGCAC	CTTCCTACGA	AGAGTTGTCA	300
20	AACTCTCAGG	AACTGCTGGA	AACCGGTATC	AAAGTTATCG	ACCTGATGTG	350
	TCCGTTCCGC	AAGGCGGTA	AAGTCGGTCT	GTTTCGGTGGT	GCGGGTGTAG	400
	GTAAAACCGT	AAACATGATG	GAGCTTATTC	GTAACATCGC	GATCGAGCAC	450
	TCCGGTTACT	CAGTGTTTGC	GGGCGTAGGG	GAACGTACTC	GTGAGGGTAA	500
	CGACTTCTAC	CACGAAATGA	CCGACTCCAA	CGTTATCGAT	AAAGTATCCC	550
25	TGGTGTATGG	CCAGATGAAC	GAGCCGCCGG	GAAACCGTCT	GCGCGTTGCA	600
	CTGACCGGCC	TGACCATGGC	GGAGAAATTC	CGTGACGAAG	GTCGTGACGT	650
	ACTGCTGTTC	GTCGATAACA	TCTATCGTTA	CACCCTGGCC	GGTACGGAAG	700
	TATCCGCACT	AGTGGGCCGT	ATGCCTTCCG	CAGTAGGTTA	CCAGCCGACT	750
	CTGGCGGAAG	AGATGGGCGT	TCTGCAGGAA	CGTATCACCT	CCACCAAAC	800
30	CGGTTCTATC	ACCTCCGTA				819

2) INFORMATION FOR SEQ ID NO: 761

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 812 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

45

(A) ORGANISM: *Yersinia pestis*  
 (B) STRAIN: KIM D27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 761

50	ACCAAAAGTG	TACAACGCC	TTGAGGTAGA	AGGTACAAC	GAAAAGTTAG	50
	TGCTGGAAGT	TCAGCAACAG	TTGGGCGGTG	GTGTTGTTCG	TTGTATCGCA	100
	ATGGGCTCTT	CCGATGGTCT	GAGCCGTGGG	TTGAAAGTAA	CCAACCTAGA	150
	ACACCCGATC	GAAGTACCGG	TTGGTAAAGC	GACCCTTGGC	CGTATCATGA	200
	ACGTATTGGG	TGAACCAATC	GACATGAAAG	GTCCTATCGG	TGAAGAAGAG	250
55	CGTTGGGCAA	TCCATCGCGA	AGCGCCTTCT	TATGAAGAGC	TTGCTAGCTC	300
	ACAAGATCTG	TTAGAAACCG	GTATCAAGGT	TATGGACCTG	ATTTGTCCGT	350
	TTGCTAAGGG	CGGTAAAGTC	GGTCTGTTTC	GTGGTGCGGG	TGTAGGTAAA	400
	ACAGTAAACA	TGATGGAGCT	GATCCGTAAC	ATCGCGATCG	AGCACTCTGG	450
	GTATTCTGTA	TTTGCCGGTG	TAGGTGAGCG	TACCCGTGAG	GGTAATGACT	500
60	TCTACCATGA	AATGACTGAC	TCCAACGTTT	TGGACAAAGT	ATCCTTGGTT	550

TACGGCCAGA TGAATGAGCC ACCAGGTAAC CGTCTACGCG TTGCACTGAC 600  
 CGGCCTGACC ATGGCCGAGA AATTCCGTGA TGAAGGTCGT GACGTACTGC 650  
 TGTTTATCGA TAATATCTAT CGTTATACCC TAGCTGGTAC GGAAGTATCC 700  
 GCATTGCTGG GTCGTATGCC ATCAGCGGTA GGTATCAGC CAACACTGGC 750  
 5 TGAAGAGATG GGTGTGTTGC AGGAACGTAT TACTTCCACT AAGACGGGTT 800  
 CAATCACCTC TG 812

10 2) INFORMATION FOR SEQ ID NO: 762

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 408 bases  
 (B) TYPE: Nucleic acid  
 15 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

20 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Burkholderia mallei*  
 (B) STRAIN: GB8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 762

25 CGACGTGATC GAGCCCTTCT TCGTCGACGT GATGCGCTCT TGCAGCTTGC 50  
 CCATTTCTTC AGCCAGCGTC GGCTGATAGC CCACTGCCGA CGGCATACGG 100  
 CCGAGCAGCG CCGACACTTC GGTACCGGCC AGCGTGAAAC GGTAGATGTT 150  
 GTCGACGAAG AACAGCACGT CGAGGCCTTC GTCACGGAAG TGCTCGGCCA 200  
 30 TCGTGAGGCC CGTCAGCGCC ACGCGCAGAC GGTGCCCCGG CGGCTCGTTC 250  
 ATCTGGCCGT ACACCAGCGC GACCTTGTCG AGAACGTTTC AGTCCTTCAT 300  
 TTCGTGGTAG AAGTCGTTCC CTTACGGGT ACGCTCGCCC ACGCCCGCGA 350  
 ACACGGAGTA ACCGCCGTGC TCCTTCGCGA TGTTGTTGAT GAGCTCCATC 400  
 ATGTTGAC 408  
 35

2) INFORMATION FOR SEQ ID NO: 763

40 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 45 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Clostridium sordellii*  
 50 (B) STRAIN: ATCC 9714

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 763

55 GAACCTATAA ACAACATAGC TACTCAACAT GGTGGTATAT CAGTATTCGC 50  
 AGGTGTTGGA GAGAGAACAA GAGAAGGTAA CGACCTTTAC GGAGAAATGA 100  
 GTGAGTCTGG AGTTATAAAC AAGACAGCTC TAGTATTCGG ACAAATGAAT 150  
 GAGCCACCTG GAGCAAGAAT GAGAGTTGCT TTAAGTGGTC TTACAATGGC 200  
 TGAATATTTT AGAGATCAAG AAGGACAAGA CGTTTTATTA TTCGTAGATA 250  
 ATATATTCCG TTTCACTCAA GCAGGATCTG AGGTTTCTGC ACTTCTTGGA 300  
 60 CGTACTCCAT CAGCAGTTGG ATACCAACCA ACATTAGCTA CAGAGATGGG 350

TAGATTACAA GAGAGAATAA CATCTACAAA TAAAGGGTCT ATAACATCAG 400

5 2) INFORMATION FOR SEQ ID NO: 764

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 405 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium novyi*  
(B) STRAIN: ATCC 19402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 764

TTAATTCAAG	AATTAATCAA	CAATATAGCG	AAGGAACACG	GTGGATTATC	50
TGTATTTACA	GGAGTTGGAG	AAAGAACAAG	AGAAGGTAAT	GACCTTTACT	100
ATGAAATGAA	AGAATCTGGT	GTTATAAATA	AAACAGCACT	AGTATTTGGT	150
CAAATGAATG	AGCCACCTGG	AGCAAGAATG	AGAGTTGCTC	TTACAGGACT	200
TACTATGGCA	GAATATTTCA	GGGATCAAGG	ACAAAACGTA	CTTTTATTCA	250
TAGACAACAT	ATTTAGATTTC	ACTCAAGCAG	GTTCAGAGGT	GTCAGCTTTA	300
CTTGGAAGAA	TACCTAGTGC	CGTTGGATAC	CAACCAACAC	TAGCAACAGA	350
AATGGGTGCG	CTTCAAGAAA	GAATAACATC	TACAAAGCAT	GGTTCATATA	400
CATCA					405

2) INFORMATION FOR SEQ ID NO: 765

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 393 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium botulinum*  
(B) STRAIN: 20:3.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 765

AATTAATAAA	TAACATAGCT	AAAGAACATG	GTGGATTATC	AGTGTTTACT	50
GGAGTTGGAG	AAAGATCTAG	AGAAGGTAAT	GATCTTTATC	ATGAAATGAG	100
AGAATCAGGC	GTTATAGATA	AAACCGCATT	AGTATTTGGT	CAAATGAATG	150
AGCCACCTGG	TGCTAGAATG	AGAGTTGCAT	TAACAGGGTT	AACTATGGCT	200
GAATATTTTA	GAGATAAAGG	TCAAGATGTA	TTACTATTTA	TAGATAATAT	250
ATTTAGATTT	ACTCAAGCTG	GATCAGAAGT	TTCAGCATT	CTTGGAAGAA	300
TACCATCAGC	AGTTGGTTAC	CAACCAACTT	TAGCAACTGA	AATGGGTGCA	350
CTTCAAGAGA	GAATTACATC	AACTAAGAAT	GGTTCATATA	CAT	393

60 2) INFORMATION FOR SEQ ID NO: 766

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium histolyticum*  
 (B) STRAIN: ATCC 19401

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 766

15	AGGAACTTAT	AAATAATATT	GCAAAACAAT	ATGGAGGTCT	ATCTGTATTT	50
	ACAGGTGTTG	GTGAAAGAAC	AAGAGAAGGT	AATGACCTAT	ATAATGAAAT	100
	GAAAGATTCT	GGGGTTATAG	AGAAACTGC	ACTAGTATTT	GGACAGATGA	150
	ATGAACCACC	AGGAGCGAGA	ATGAGAGTAG	CATTGACAGG	ACTTACTATG	200
20	GCAGAAATATT	TTAGAGATCA	AGGGCAAGAT	GTACTTTTAT	TTATAGATAA	250
	TATATTTAGA	TTTACGCAGG	CTGGTTCTGA	AGTTTCTGCA	TTGTTAGGAA	300
	GAATTC CAAG	TGCAGTTGGA	TATCAACCAA	CCCTTGCAAC	TGAAATGGGT	350
	GCATTACAAG	AAAGAATAAC	ATCCACAAAA	AATGGATCAA	TTACTTCAG	399

## 2) INFORMATION FOR SEQ ID NO: 767

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 390 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Peptostreptococcus prevotii*  
 (B) STRAIN: ATCC 9321

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 767

45	CATGATGGAA	CTGATCCGTA	ACATCGCCAT	CGAGCACAGC	GGTTATTCCG	50
	TGTTTCGCCG	TGTGGGTGAG	CGTACTCGTG	AGGGTAACGA	CTTCTACCAC	100
	GAGATGAAGG	ACTCCAACGT	TCTGGACAAA	GTGGCACTGG	TCTACGGTCA	150
	GATGAACGAG	CCGCCGGGTA	ACCGTCTGCG	CGTAGCACTG	ACTGGCCTGA	200
	CCATGGCCGA	GAAGTTCCGT	GACGAAGGTA	ACGACGTTCT	GTTGTTCGTC	250
	GACAACATCT	ACCGTTACAC	CCTGGCCGGT	ACTGAAGTAT	CCGCACTGCT	300
	GGGCCGTATG	CCTTCGGCAG	TAGGTTACCA	GCCGACCCTG	GCTGAAGAGA	350
50	TGGGCGTTCT	GCAAGAACGT	ATCACTTCGA	CCAAGGAAGG		390

## 2) INFORMATION FOR SEQ ID NO: 768

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 953 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Absidia corymbifera*  
(B) STRAIN: ATCC 46775

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 768

```

10 AGGTCTTGTT CGTGGCAAAA GGTCATTGAC ACTGGTGCTC CTATCACCAT      50
   TCCTGTTGGT AACGAAGTCC TTGGTCGTAT CATTAACGTC ATTGGTGAGC      100
   CCATTGATGA GCGTGGTCCY ATCAAGTCCA AGGCCACTCG TGCTATCCAC      150
   GCTGATGCTC CCGAGTTCGT TGATCAATCC CCCACTCCCG AGATTCTCGA      200
   GACTGGTATC AAGGTTGTCT ATTTGCTTGC TCCCTATGCT CGTGGTGGTA      250
15 AGATTGGTCT TTTCGGTGGT GCCGGTGTCT GCAAGACTGT CTTGATTCAA      300
   GAATTGATCA ACAACATTGC CAAGGCTCAC GGTGGTTACT CTATCTTCTG      350
   TGGTGTCTGGT GAACGTACTC GTGAAGGCAA CGATTGTGAC CACGAAATGA      400
   TTTCCACTGG TGTCATCAAG CTTGAAGGTG AATCCAAGTG TGCTCTTGTC      450
   TTTGGTCAAA TGAACGAACC CCCCAGGAGCT CGTGCCCGTG TTGCCCTGAC      500
20 TGGTTTGACC ATTGCCGAAT ACTTCCGTGA TGAGGAAGGT CAAGATGTGT      550
   TGCTCTTCAT TGACAACATT TTCCTGTTCA CTCAAGCCGG TTCTGAAGTG      600
   TCCGCTTTGC TTGGTCTGAT TCCCTCTGCT GTCGGTTACC AACCCACTCT      650
   YTCCACTGAT ATGGGTGGTA TGCAAGAGCG TATTACTACC ACCAAGAACG      700
   GTTCCATTAC CTCCGTGCAA GCTGTCTACG TCCCTGCTGA CGATTTGACT      750
25 GATCCTGCTC CTGCTACTAC TTTTGCTCAC TTGGACGCCA CCACTGTGTT      800
   GTCTCGTTCC ATTGCTGAGT TGGGTATCTA CCCTGCTGTC GATCCCCTTG      850
   ACTCCAAGTC CCGTATCTTG GATCCCCGTA TCGTTGGTGA TGAGCACTAC      900
   TCTGTTGCCA CTGGTGTCCA ACAAATCCTT CAAAACACTACA AGTCGCTTCA      950
   AGA
30

```

2) INFORMATION FOR SEQ ID NO: 769

- 35 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 1343 bases  
     (B) TYPE: Nucleic acid  
     (C) STRANDEDNESS: Double  
     (D) TOPOLOGY: Linear

40 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 45 (A) ORGANISM: *Alternaria alternata*  
(B) STRAIN: ATCC 62099

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 769

```

50 CGCGGTTCCA AGGCCACCGA CACCGGTGCC CCCATCAAGA TTCCCGTTGG      50
   TCACGGTACC CTTGGTCGTA TCATGAACGT CACTGGTGAC CCCATTGACG      100
   AGCGTGGTCC CATCAAGGCC ACCAAGTACG CTCCCATCCA CGCCGACCCC      150
   CCGGAGTTCA CCGAGCAATC CACCTCCGCT GAGGTCCTCG TTACCGGTAT      200
   CAAGGTTGTC GACCTGTTGG CTCCTTACGC TCGTGGTGGA AAGATTGGTC      250
   TCTTCGGAGG TGCTGGTGTC GGAAAGACTG TCTTCATTCA GGAGCTGATT      300
55 GTAAGGAGAC ACACTGTCTA CTGGCTGAGC ATTAGCTAAC GGCAGGCAGA      350
   ACAACATCGC CAAGGCCAC GGTGGTTTCT CTGTCTTCAC TGGTGTCTGGT      400
   GAGCGTACCC GTGAGGGTAA CGATCTGTAC CACGAGATGC AGGAGACTTC      450
   CGTCATTGAG CTTGACGGTG ACTCCAAGGT CGCCCTCGTC TTCGGTCAGA      500
   TGAACGAGCC CCGGGGTGCC CGTGCCCGTG TCGCTCTTAC TGGTCTTACT      550
60 GTTGCTGAGT AAGTCTTGAA TTCACTGTGT TGACAACGTC GTGGCTAATG      600

```



	GGAAAAAAGA	TACTTCCGTG	ACGAAGAGGG	ACAGGATGTG	CTTCTCTTCA	650
	TCGACAGTAA	GTGCTTGTA	GAAGTGCCTG	TGAGACATAC	ACTGACTTCG	700
	GCAATAGACA	TTTTCCGTTT	CACCCAGGCC	GGTCCGAGG	TGTCCGCTCT	750
	TCTTGGTTCG	ATCCCCCTCTG	CCGTCGGTTA	CCAGCCCACC	CTCGCCATTG	800
5	ACATGGGTGT	CATGCAGGAG	CGTATTACCA	CCACCACCAA	GGGTTCATC	850
	ACCTCCGTCC	AGGCCGTCTA	CGTGCCCGCT	GACGATTTGA	CTGACCCTGC	900
	CCCCGCCACC	ACCTTCGCCC	ATTTGGACGC	CACCACTGTC	TTGTCCCGTG	950
	GTATCTCCGA	GTTGGGTATC	TACCCCGCCG	TCGACCCTCT	TGACTCCAAG	1000
	TCCCGTATGT	TGGACCCCCG	TGTCATTGGT	CAGGACCACT	ACGACACCGC	1050
10	CACCCGCGTT	CAGCAGATT	TCCAGGAGTA	CAAGTCGCTC	CAGGATATCA	1100
	TTGCCATTCT	CGGTATGGAC	GAGTTGTCGG	AAGCTGACAA	GCTTACCGTC	1150
	GAGCGTGCCC	GTAAGATCCA	GCGTTTCTTG	AGCCAGCCTT	TCGCTGTCGC	1200
	CCAGGTCTTC	ACTGGTATTG	AGGGCAAGCT	TGTCGACCTC	AAGGACACCA	1250
	TCCGATCATT	CAAGGCTATC	TTGACTGGTG	AGGGTGACGA	CCTTCCCGAG	1300
15	GGTGAGTCTC	GACTATCTCC	GCATTCATAG	CGTATAACTG	ACA	1343

## 2) INFORMATION FOR SEQ ID NO: 770

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 480 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - 25 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- 30 (A) ORGANISM: *Aspergillus flavus*
  - (B) STRAIN: ATCC 26947
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 770

35	TTCAGGAGTT	GATTGTATGT	TCACCTGCAA	CATAAGACTT	CCCATTCTCC	50
	ACTCTTTTCT	AACTCTTCAC	AGAACAACAT	TGCCAAGGCT	CACGGTGGTT	100
	ACTCTGTCTT	CACTGGTGTC	GGTGAGCGTA	CCCGTGAGGG	TAACGATCTG	150
	TACCACGAAA	TGCAGGAGAC	TGGTGTCATT	CAGCTCGAGG	GTGAATCTAA	200
	GGTCGCCCTT	GTCTTCGGTC	AGATGAACGA	GCCCCCAGGT	GCCCCGTGCCC	250
40	GTGTCGCCCT	TACCGGTCTG	ACCATCGCCG	AGTACTTCCG	TGACGAGGAA	300
	GGTCAGGATG	TGCTGCTCTT	CATTGACAAC	ATTTTCCGTT	TCACCCAGGC	350
	CGGTTCTGAG	GTGTCTGCCC	TTCTTGGTG	TATCCCCCTC	GCTGTCGGTT	400
	ACCAGCCCAC	TCTGGCCGTC	GACATGGGTG	GTATGCAGGA	GCGTATTACC	450
45	ACCACCACCA	AGGGTTCCAT	TACCTCCGTC			480

## 2) INFORMATION FOR SEQ ID NO: 771

- 50 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1174 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - 55 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- 60 (A) ORGANISM: *Mucor circinelloides*
  - (B) STRAIN: ATCC 38592

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 771

```

5   CTCGAACAAG AYAACCTTGCC YGCCATTTTG AACGCCCTTG AAGTCAAGGA      50
   YCACTCTGGT GGACGTCTCG TYCTCGAAGT CTCTCAACAT TTGGGTGAGA      100
   ACACGTGTCG TACTATTGCT ATGGATGGTA CTGAAGGTAA GTTATGTYCA      150
   TCCCANNGGA TACAGTCARA CAGMAATGTC TAGTGGTTAT AGCAGYAGCA      200
   GATGATTGAC CAATATGTTA GGTCTTGTCC GGTGGTCAAA AGGTTGTTGA      250
   CACTGGTGCT CCCATTACCA TCCCCGTTGG TAAGGAAGTC CTTGGTCGTA      300
10  TCATCAACGT TATTGGTGAA CCCATTGATG AACGTGGTCC CATTGACGCC      350
   AAGACTCACC GTCCTATTCA CGCTGAAGCT CCCGAATTCG TTGATCAATC      400
   CCCCCTCCC GAAATCCTCG AGACTGGTAT CAAGGTYGTC GATTTGTTGG      450
   CTCCTTATGC TCGTGGTGGT AAGATTGGTC TCTTCGGTGG TGCTGGTGTC      500
   GGTAAAGACTG TCTTGATTCA AGAATTGATT AACAACATYG CCAAGGCTCA      550
15  CGGTGGTTAC TCTATCTTCT GTGGTGTCCG TGAACGTACT CGTGAGGGTA      600
   ACGATTTGTA CCATGAAATG ATTGAAACCG GTGTCATTCA ATTGGAAGGC      650
   GAGTCCAAGT GTGCTCTCGT CTTTGGTCAA ATGAACGAAC CCCAGGTGC      700
   TCGTGCCCGT GTCGCTTTGA CTGGTTTGAC TATTGCTGAA TACTTCCGTG      750
   ATGATGAGGG TCAAGATGTC TTGCTTTTCA TTGATAACAT TTTCAGATTC      800
20  ACTCAAGCTG GTTCTGAGGT ATCTGCCCTT TTGGGTCGTA TTCCTTCCGC      850
   TGTCGGTTAC CAACCCACTC TTTCCACYGA CATGGGTGGT ATGCAAGAGC      900
   GTATTACTAC CACCAAGAAC GGTTCCATTA CCTCCGTCCA AGCTGTCTAC      950
   GTCCCTGCTG ATGATTTGAC CGATCCTGCT CCTGCCACCA CTTTGTGCTCA     1000
   CTTGGATGCC ACCACTGTCT TGTCTCGTTC CATCGCTGAA TTGGGTATCT     1050
25  AYCCCGCTGT CGATCCTCTT GATTCCAAGT CTCGTATCCT CGATCCCCGT     1100
   ATTGTCGGTG ATGAGCACTA CAAGGTTGCC ACTGAAGTTC AACAAATCCT     1150
   CCAAACACTAC AAGTCTCTCC AAGA                                     1174

```

30

## 2) INFORMATION FOR SEQ ID NO: 772

## (i) SEQUENCE CHARACTERISTICS:

```

35  (A) LENGTH: 467 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
   (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

40

## (vi) ORIGINAL SOURCE:

```

   (A) ORGANISM: Piedraia hortai
   (B) STRAIN: ATCC 24292

```

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 772

```

   AGGAGCTTAT CAACAACATC GCCAAGGCTC ACGGTGGTTA CTCCGTCTTC      50
   ACTGGTGTCG GTGAGCGTAC TCGTGAGGGT AACGATTTGT ACCACGAGAT      100
   GCAAGAGACT TCCGTCAATC AGCTTGACSG CGAGTCCAAG GTCGCTCTCG      150
50  TGTTCCGGTCA GATGAACGAG CCCCCGGGTG CCCGTGCCCC TGTTGCCCTG      200
   ACTGGTCTTA CCATCGCTGA GTACTTCCGT GATGCCGAGG GTCAGGATGG      250
   TAAGTTCTAT AACTCTTGTC GCAAAGGTTT CATTCTGGTC GCTAACTTGC      300
   TCAGTGCTCC TGTTTCATCGA CAACATTTTC CGTTTCACCC AGGCTGGTAT      350
   GGAGGTGTCT GCCCTCCTCG GTCGTATTCC TTCTGCCGTC GGTTACCAAC      400
55  CTACTCTCGC CGTCGACATG GGTGGTATGC AAGAGCGTAT TACCACTACC      450
   AAGAAGGGAT CCATTAC                                     467

```

## 60 2) INFORMATION FOR SEQ ID NO: 773

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 578 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Pseudallescheria boydii*  
 (B) STRAIN: ATCC 44331

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 773

```

15 TTGCCAAGGC CCACGGTGGT TACTCTGTCT TCACTGGTGT CGGTGAGCGT      50
   ACCCGTGAGG GTAACGATCT GTACCACGAA ATGCAGGAAA CCTCCGTCAT      100
   TCAGCTTGAT GGCAGAGTCCA AGGTCGCGCT TGTCTTCGGT CAGATGAACG      150
   AGCCCCCTGG TGCCCGTGCT CGTGTCGCTC TTACTGGTCT TACCGTTGCC      200
20 GAGTACTTCC GTGATGAGGA GGGTCAGGAT GGTAAGTTAT ATCGTTTTTA      250
   TTATCTTCTT TGCCACCACC CCTCTACGAA TCCATGCCTC CGTTGGTGAA      300
   GGCATCGTTT GTAGGGCGGG TCGGAGTTTG CGGCAATTTC TGCCGTCGGC      350
   TTGAAGCCGC GGATGCCCGA TGTTTGACGC GTATCGATGC TAACAACAAT      400
   GACAACAGTG CTTCTCTTCA TTGACAACAT TTTCCGATTC ACCCAGGCCG      450
25 GTTCCGAGGT GTCTGCCCTT TCGGTCGTA TTCCCTCTGC CGTCGGTTAC      500
   CAGCCCACTC TTGCCGTAGA CATGGGTGCC ATGCAGGAAC GTATTACCAC      550
   CACCAAGAAG GGTTCGATTA CCTCCGTC      578

```

## 2) INFORMATION FOR SEQ ID NO: 774

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1123 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Rhizopus oryzae*  
 (B) STRAIN: ATCC 56015

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 774

```

50 AACTTACCYG CTATCTTGAA CGCTCTCGAA GTCCAAGATC ACTCTGGTGG      50
   ACGTCTTGTC CTTGAAGTTC GCTCAACACT TGGGTGAAAA TACTGTCCGT      100
   ACTATTGCTA TGGATGGTAC TGAAGGTAAG CTATACTATA ACCGTKTTAT      150
   CCGAGTATGA TATTAAGTTG AAAAAAGGTC TCGTYCGTGG TCAAAAAGGTT      200
   ATTGACACTG GTGCTCCCAT TACCATTCCCT GYTGGTAAGG AAGTTCTCGG      250
   TCGTATCATT AACGTCATTG GTGAACCCAT CGATGAACGT GGTCCCTATCA      300
   ACGCCAAGAG CCAACGTCCC ATTCACGCCG AAGCTCCCGA ATTTCGTTGAC      350
   CAATCTCCTA CTCCCGAAAT TCTTGAAACT GGTATCAAGG TTGTCGACTT      400
55 GTTGGCTCCT TATGCTCGTG GTGGTAAGAT TGGTCTTTTC GGTGGTGCTG      450
   GTGTCGGTAA GACTGTGTTG ATTCAAGAAT TGATTAACAA CATCGCCAAG      500
   GCTCACGGTG GTTACTCTAT TTTCTGTGGT GTCGGTGAAC GTACTCGTGA      550
   AGGTAACGAT CTTTACCACG AAATGATTGA AACTGGTGTC ATCAAGCTCG      600
   ATGGTGACTC CAAGTGTGCT CTTGTCTTTG GTCAAATGAA CGAACCCCCA      650
60 GGAGCTCGTG CCCGTGTTGC CTTGACTGGT TTGACCATTG CTGAATACTT      700

```

CCGTGATGAT	GAAGGTCAAG	ATGTGTTGCT	TTTCATTGAT	AACATTTTCC	750
GTTTCACCCA	AGCTGGTTCW	GAAGTATCTG	CCCTTTTGGG	TCGTATTCCC	800
TCCGCTGTCG	GTTACCAACC	CACTCTTTCT	ACTGATATGG	GTGGTATGCA	850
AGAACGTATT	ACAACCACCA	AGAACGGTTC	CATTACATCC	GTCCAAGCTG	900
5 TCTACGTCCC	TGCTGATGAT	TTGACCGATC	CTGCTCCCGC	CACCACTTTT	950
GCTCACTTGG	ATGCCACCAC	TGTCTTGTCT	CGTTCCATTG	CCGAATTGGG	1000
TATTTACCCT	GCCGTCGATC	CTCTTGAYTC	CAAGTCTCGT	ATCTTGGATC	1050
CTCGTATCGT	TGGTGACGAA	CACTACAAGG	TCGCTACCGA	AGTTCAACAA	1100
ATCCTTCAAA	ACTACAAGTC	TCT			1123

10

## 2) INFORMATION FOR SEQ ID NO: 775

- 15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 477 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 20 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Scopulariopsis koningii*  
 25 (B) STRAIN: ATCC 38745

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 775

ATTCAGGAGC	TCATCAACAA	CATTGCCAAG	GCTCACGGTG	GTTACTCTGT	50
30 GTTCACTGGT	GTCGGTGAGC	GTACCCGTGA	GGGTAACGAT	CTGTACCACG	100
AAATGCAGGA	GACTTCGGTC	ATTCAGCTCG	AGGGCGAGTC	CAAGGTCGCG	150
CTTGTGTTTCG	GTCAGATGAA	CGAGCCCCCC	GGTGCCCGTG	CCCGTGTCGC	200
CCTTACCGGT	CTGACCGTTG	CCGAGTACTT	CCGTGACGAG	GAGGGCCAGG	250
ATGGTGAGTA	ACCGACGAAG	TCTGAGATCT	TGTCGGGCAT	TATTCTAACG	300
35 ACAACTAGTG	CTTCTCTTCA	TCGACAACAT	TTTCCGCTTC	ACCCAGGCCG	350
GTTCCGAGGT	GTCCGCGCTT	CTCGGCCGTA	TCCCCTCTGC	CGTCGGTTAC	400
CAGCCCACCC	TGGCCGTCGA	CATGGGAGGT	ATGCAGGAGC	GTATTACCAC	450
GACTCAGAAG	GGCTCGATTA	CCTCGGT			477

40

## 2) INFORMATION FOR SEQ ID NO: 776

- 45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 610 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 50 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Trichophyton mentagrophytes*  
 (B) STRAIN: ATCC 8125

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 776

TCCGGAGTTG	ATTGTAAGTC	ATTTGAAACC	CAGCCCCAAG	AAACAGAAGC	50
TAGGCGAAAA	TTGGACAATT	GAGCAATTTA	GCCATTGGAG	AAAAGAAATT	100
60 TCGAGTATTA	ATTGTTTTTA	TAGAACAACA	TTGCCAAGGC	TCACGGTGGT	150

TACTCTGTCT TCACTGGTGT CGGAGAGCGT ACCCGTGAAG GAAACGATCT 200  
 CTACCATGAA ATGCAGGAGA CCCGTGTCAT TCAGCTTGAT GGCGAGTCCA 250  
 AGGTGCCCCCT GGTCTTCGGC CAGATGAACG AGCCCCCAGG TGCCCGTGCC 300  
 CGTGTTGCTC TTACTGGTTT GACCATTGCT GAGTACTTCC GTGATGAGGA 350  
 5 AGGTCAAGAC GGTGAGTTTY TTATGGATAA AARATTTTTT TTTTTTTTTT 400  
 TTTTTTMAAR AAATTCATGT TCTAACAAAG TATATCCTAG TGCTTCTCTT 450  
 CATCGACAAC ATTTTCCGTT TCACTCAGGC TGGTTCCGAA GTGTCTGCCC 500  
 TGCTTGGTTCG TATCCCATCT GCCGTCGGTT ACCAACCAC TCTTGCCGTC 550  
 GACATGGGTG GTATGCAGGA ACGTATTACC ACCACCAAGA AGGGATCCAT 600  
 10 TACCTCCGTC 610

2) INFORMATION FOR SEQ ID NO: 777

- 15 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 593 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - 20 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- 25 (A) ORGANISM: *Trichophyton tonsurans*
  - (B) STRAIN: ATCC 56185
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 777

30 GGAGTTGATT GTAAGTCATT TGAAACCCAG CCCCAAGAAA CAGAAGCTAG 50  
 GTGAAAATTG GACAATTGAA CAATTTAGCC CTTGGAGAAA AGAAATTTTCG 100  
 AGTATTAATT ATTTTATAG AACAACATTG CCAAGGCTCA CGGTGGTTAC 150  
 TCTGTCTTCA CTGGTGTCTCG AGAGCGTACC CGTGAAGGAA ACGATCTCTA 200  
 CCATGAAATG CAGGAGACCC GTGTCAATCA GCTTGATGGC GAGTCCAAGG 250  
 35 TCGCCCTGGT CTTCGGCCAG ATGAACGAGC CCCAGGTGC CCGTGCCCGT 300  
 GTTGCTCTTA CTGGTTTGAC CATTGCTGAG TACTTCCGTG ATGAGGAAGG 350  
 TCAAGACGGT GAGTTTCTTA TGGATGAAAG ATTTTTTTTT TTCAAGAAAT 400  
 TCATGTTCTA ACAAAGTGTA TTCTAGTGCT TCTCTTCATC GACAACATTT 450  
 TCCGTTTCAC TCAGGCTGGT TCCGAAGTGT CTGCCCTGCT TGGTCGTATC 500  
 40 CCATCTGCCG TCGGTTACCA ACCCACTCTT GCCGTCGACA TGGGTGGTAT 550  
 GCAGGAACGT ATTACCACCA CCAAGAAGGG ATCCATTACC TCC 593

45 2) INFORMATION FOR SEQ ID NO: 778

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1141 bases
  - (B) TYPE: Nucleic acid
  - 50 (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- 55 (A) ORGANISM: *Trichosporon cutaneum*
  - (B) STRAIN: ATCC 62965
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 778
- 60

	CCGTGGTCAA	GAAGTTATTG	ACACTGGTGC	CCCAATTACC	ATTCCTGTTG	50
	GTCGTGGTAC	TCTTGGTAGA	ATTATCAACG	TCATTGGTGA	ACCAATTGAC	100
	GAACGTGGCC	CTATCAAGGC	TTCTAAGTAT	GCTCCTATCC	ATACTGAACC	150
5	ACCAACCTTT	GCTGAACAAT	CTACTTCTGC	TGAAGTTCTT	GAAACCGGTA	200
	TCAAGGTTGT	CGATCTTCTT	GCTCCTTACG	CCCGTGGTGG	TAAGATTGGT	250
	CTTTTCGGTG	GTGCTGGTGT	CGGTAAGACT	GTCTTCATTG	AAGAACTTAT	300
	TAACAACATT	GCCAAGGCTC	ACGGTGGTTT	CTCTGTCTTC	ACTGGTGTCG	350
	GTGAAAGAAC	CCGTGAAGGT	AACGATCTTT	ACCGTGAAAT	GAAGGAAACT	400
	GGTGTCATCA	ACCTCGAAGG	TGACTCCAAG	GTCGCTCTCG	TTTTTCGGTCA	450
10	AATGAACGAA	CCTCCAGGTG	CCCGTGCCCG	TGTCGCTTTG	ACTGGTCTTA	500
	CCATTGCCGA	ATACTTCCGT	GATGAAGAAG	GACAAGATGT	CTTGCTTTTC	550
	GTTGACAACA	TTTTTCAGATT	CACCCAAGCC	GGTTCTGAAG	TCTCTGCTCT	600
	TTTGGGTTCG	ATTCCATCTG	CCGTCGGTTA	CCAACCTACC	CTTGCTACCG	650
	ATATGGGTGC	CCTCCAAGAA	CGTATTACCA	CCACCCAATA	GGGTTCGGTC	700
15	ACATCTGTCC	AAGCCGTCTA	TGTCCCAGCA	GACGATTTGA	CTGATCCTGC	750
	CCCAGCCACC	ACTTTCGCTC	ACTTGGACGC	CACCACTGTC	TTGTCTCGTT	800
	CCATTTCCGA	ATTGGGTATC	TACCCAGCTG	TCGATCCTCT	CGATTCCAAG	850
	TCTCGTCTTT	TGGATCCTGA	AGTTATTGGA	CACGAACACT	ACGAAGTTGC	900
	CACTCAAGTT	CAACAAACCC	TCCAAGCTTA	CAAGTCTCTC	CAAGATATCA	950
20	TTGCCATTTT	GGGTATGGAT	GAATTGTCTG	AAGCTGATAA	GCTTACTGTC	1000
	GAACGTGCCC	GTAAGATCCA	AAGATTCCTT	TCCCAACCAT	TCGCTGTTGC	1050
	CGAAGTTTTT	ACTGGTATCG	AAGGTCGTCT	CGTTCCATTG	AAGGAAACCG	1100
	TCAGATCTTT	CAAGGAAATC	CTTGAAGGTA	AGTACGATCA	C	1141

25

## 2) INFORMATION FOR SEQ ID NO: 779

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 1093 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cladophialophora carrionii*  
 (B) STRAIN: ATCC 22864

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 779

	GCTGAAGGCC	GAACGTGAGC	GTGGTATCAC	CATCGATATC	GCGCTCTGGA	50
	AGTTCGAGAC	TCCCAAGTAC	TTCGTCACCG	TCATCGATGC	CCCTGGTCAT	100
45	CGTGACTTCA	TCAAGAACAT	GATCACTGGT	ACCTCCCAGG	CTGATTGTGC	150
	TATTCTCATC	ATTGCCGCTG	GTACTGGTGA	GTTCGAGGCC	GGTATCTCCA	200
	AGGATGGCCA	GACCCGTGAG	CATGCTCTGC	TCGCCTACAC	CCTGGGCGTG	250
	AAGCAGCTTA	TCGTCGCCAT	CAACAAGATG	GACACCACCA	AATGGTCTGA	300
	GGATCGTTTC	AACGAAATCA	TCAAGGAGAC	TTCCAAC TTC	ATCAAGAAGG	350
50	TCGGATACAA	CCCCAAGTCC	GTTCCATTCT	TGCCCCATCTC	CGGCTTCAAC	400
	GGTGACAACA	TGATCGACGT	CTCCACCAAT	GCCCCCTGGT	ACAAGGGCTG	450
	GGAAAAGGAG	TCCAAGGCTG	GCAAGGCCAC	CGGCAAGACC	CTCCTTGAGG	500
	CTATCGACTC	CATCGACCCT	CCTGCTCGTC	CCACCGACAA	GCCTCTCCGT	550
	CTCCCACTCC	AGGATGTCTA	CAAGATTTCT	GGTATCGGCA	CGGTGCCCCG	600
55	CGGTCTGTGT	GAGACTGGTA	CCATCAAGGC	CGGTATGGTC	GTCACCTTTG	650
	CCCCCGCCAA	CGTCACCACT	GAAGTCAAGT	CCGTCGAAAT	GCACCACGAA	700
	CAGCTCGCCG	AGGGCGTTCC	GGGTGACAAC	GTCGGCTTCA	ACGTCAAGAA	750
	CGTCTCCGTG	AAGGAGGTTT	GTCGTGGAAA	CGTTGCTGGT	GACTCCAAGA	800
	ACGACCCCCC	CAAGGGTGCC	GACTCCTTCA	ACGCCCAGGT	CATCGTCCTC	850
60	AACCACCTTG	GTCAGGTCGG	TGCTGGCTAC	GCCCCGGTCT	TGGATTGCCA	900

CACTGCCCAC	ATTGCCTGCA	AGTTCTCTGA	GCTCCTCGAG	AAGATCGATC	950
GTCGTACCGG	CAAGTCCATG	GAAAACAACC	CCAAGTTCAT	CAAGTCTGGT	1000
GATGCTGCCA	TCGTGAAGAT	GGTTCCCAGC	AAGCCTATGT	GCGTTGAGGC	1050
TTTCACCGAC	TACCCTCCTC	TTGGTCGTTT	CGCCGTCCGT	GAC	1093

5

## 2) INFORMATION FOR SEQ ID NO: 780

## 10 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 752 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

15

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cunninghamella bertholletiae*
- (B) STRAIN: ATCC 42115

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 780

25	TACTTGTA	AA	TGGTCTCAAG	ATCGTTACAA	CGAAATTGTT	AAGGAAGTTT	50
	CTTCCTTCAT	CAAGAAGATT	GGTTACAATC	CTAAATCCGT	TCCTTTTCGT	Y	100
	CCTATCTCTG	GTTGGCACGG	TGATAACATG	TTGGAAGCTT	CTACCAACAT		150
	GCCTTGGTAC	AAGGGATGGA	CCAAGGAAAC	TAAAGCTGGT	TCTTCCACTG		200
	GTAAGACTCT	CTTAGAAGCC	ATTGACAGCA	TTGAACCTCC	TACCCGTCCT		250
	TCTGACAAGC	CTTTACGTCT	TCCTTTACAA	GATGTTTACA	AGATTGGTGG		300
30	TATTGGTACT	GTCCCTGTG	TG	GTCGTGTTGA	AACTGGTGTC	ATCAAGGCTG	350
	GTATGGTTGT	TACYTTCGCT	CCCGCTAACG	TCACCACTGA	AGTTAAGTCC		400
	GTTGAAATGC	ATCAGGAACA	ATTAGAACAA	GGTGTTCCCTG	GTGACAACGT		450
	TGGTTTCAAC	GTCAAGAACG	TTTCCGTTAA	GGATATCCGT	CGTGGTAACG		500
	TCTGTTCCGA	CTCCAAGAAC	GACCCCGCTA	AGGAATCTGC	TTCTTTCAAC		550
35	GCTCAAGTTA	TCGTCTTGAA	CCACCCTGGT	CAAATTGGTG	CTGGTTATGC		600
	CCCAGTTCTT	GACTGTGACA	CTGCTCACAT	TGCTTGTAAG	TTGCTGAAT		650
	TATTAGAAAA	GATCGATCGT	CGTTCCGGTA	AGAAACTCGA	AGATGCTCCT		700
	AAATTCGTTA	AATCTGGTGA	CTCTGCTATC	GTTAAGATGG	TTCTTTCCAA		750
	GC						752

40

## 2) INFORMATION FOR SEQ ID NO: 781

## 45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 728 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

50

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Curvularia lunata*
- (B) STRAIN: ATCC 26425

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 781

60	CAAGTGGTCT	GAGGACCGTT	ACCAGGAAAT	CATCAAGGAG	ACCTCCAAC	50
	TCATCAAGAA	GGTCGGCTAC	AACCCCAAGC	ACGTTCCCTT	CGTCCCCATC	100

	TCCGGTTTCA	ACGGAGACAA	CATGATTGAG	GCCTCCACCA	ACTGCCCCCTG	150
	GTACAAGGGT	TGGGAGAAGG	AGACCAAGGC	CAAGGCCACT	GGTAAGACCC	200
	TCCTTGAGGC	CATCGACGCC	ATCGACCCTC	CTGTCCGTCC	TACCGACAAG	250
	CCCCTCCGCC	TTCCCCCTCCA	GGATGTTTAC	AAGATTGGTG	GTATTGGCAC	300
5	GGTCCCCGTC	GGTCGTGTCTG	AGACCGGTAT	CATCAAGCCC	GGTATGGTCG	350
	TCACCTTCGC	CCCCGCTGGT	GTCACCACCG	AAGTCAAGTC	CGTCGAGATG	400
	CACCACGAGC	AGCTTACTGA	GGGTGTCCCC	GGTGACAACG	TCGGCTTCAA	450
	CGTCAAGAAC	GTCTCCGTCA	AGGAGATCCG	TCGTGGTAAC	GTGCGCGGTG	500
	ACTCCAAGAA	CGACCCCCCC	AAGGGTTGCG	AGTCCTTCAA	CGCCCAGGTC	550
10	ATCGTCCTCA	ACCACCCTGG	TCAGGTCGGT	GCCGGTTACG	CCCCAGTCCT	600
	TGACTGCCAC	ACTGCCCACA	TTGCCTGCAA	GTTCTCCGAG	CTCCTCGAGA	650
	AGATCGACCG	CCGTACCGGA	AAGTCTGTTG	AGAACTCCCC	CAAGTTCATC	700
	AAGTCCGGTG	ACGCCGCCAT	CGTCAAGA			728

15

## 2) INFORMATION FOR SEQ ID NO: 782

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1145 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Fonsecaea pedrosoi*

(B) STRAIN: ATCC 18831

30

## --(xi)-- SEQUENCE DESCRIPTION: SEQ ID NO: 782

	TGAAGTCCGA	GCGTGAGCGT	GGTATCACCA	TCGATATCGC	CCTCTGGAAG	50
	TTCGAGACTC	CCAAGTAAGG	CTCAACAGAC	ACAACAAGCA	AATGCATACT	100
35	CGCTAACCTA	TTCACCCACC	ACAGGTACAA	CGTCACCGTC	ATTGACGCCC	150
	CCGGTCACCG	TGATTTTCATC	AAGAACATGA	TCACTGGTAC	CTCCCAGGCT	200
	GACTGCGCCA	TTCTCATCAT	TGCCGCCGGT	ACTGGTGAGT	TCGAGGCCGG	250
	TATCTCCAAG	GACGGTCAGA	CCCGTGAGCA	CGCTCTTCTC	GCCTACACCC	300
	TTGGTGTCAA	GCAGCTCATC	GTGCCATCA	ACAAGATGGA	CACCACCAAG	350
40	TGGTCTGAGG	CCCGTTACCA	GGAGATCATC	AAGGAGACCT	CCGGTTTCAT	400
	CAAGAAGGTC	GGCTTCAACC	CCAAGCACGT	TCCCTTCGTG	CCCATCTCCG	450
	GTTTCAACGG	TGACAACATG	ATCGACGTCT	CCACCAACTG	CCCCTGGTAC	500
	AAGGGTTGGG	AGAAGGAGAC	CAAGGCCAAG	GCCACCGGCA	AGACCCTCCT	550
	CGAGGCCATT	GACGCCATCG	ACCCCCCAC	TCGTCCCACC	GACAAGCCCC	600
45	TCCGTCTTCC	CCTYCAGGAT	GTCTACAAGA	TCGGTGGTAT	TGGCACGGTT	650
	CCCGTCGGTC	GTGTCGAGAC	CGGTACCATC	AAGGCCGGCA	TGGTCGTCAC	700
	CTTCGCCCCC	GCTGGTGTCA	CCACTGAGGT	CAAGTCCGTC	GAGATGCACC	750
	ACGAGCAGCT	TCCCCGAGGT	CTCCCCGGTG	ACAACGTCGG	CTTCAACGTC	800
	AAGAACGTCT	CCGTCAAGGA	GATCCGTCTG	GGCAACGTCG	CCGGTGACTC	850
50	CAAGAACGAC	CCCCCAAGG	GTCGCGACAG	CTTCAACGCC	CAGGTCATCG	900
	TCCTCAACCA	CCCCGGTCAG	GTCGGCGCCG	GCTACGCGCC	CGTCCTCGAC	950
	TGCCACACTG	CTCACATTGC	TTGCAAGTTC	TCTGAGCTCC	TCGAGAAGAT	1000
	CGACCGCCGT	ACCGGCAAGT	CCATTGAGGC	CAGCCCCAAG	TTCATCAAGT	1050
	CTGGTGACGC	CGCCATCGTC	AAGATGGTTC	CCTCCAAGCC	TATGTGCGTT	1100
55	GAGGCCTTCA	CCGACTACCC	CCCTCTTGGA	CGTTTCGCCG	TCCGT	1145

## 2) INFORMATION FOR SEQ ID NO: 783

60



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1151 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Microsporium audouinii*  
 (B) STRAIN: ATCC 11347

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 783

```

15 GCTCAAGGCT GAGCGTGAGC GTGGTATCAC CATTGACATC GCCCTCTGGA      50
   AGTTCGAGAC CCCCAAGTAC ATGGTCAACG TCATCGGTAT GCTTTATCTG      100
   TTTCCCATTT ATAGTTGCGA CAAGTAACTA ATAAAAAGTA GATGCCCCCG      150
   GACACCGTGA CTTTCATCAAG AACATGATTA CTGGTACCTC CCAGGCCGAC      200
   TGCGCTATTC TCATCATTTGC TGCCGGTACT GGTGAGTTTC AGGCTGGTAT      250
20 CTCCAAGGAT GGCCAGACTC GTGAGCACGC CCTGCTCGCT TTCACCCTCG      300
   GTGTCAAGCA GCTCATCGTT GCCATCAACA AGATGGACAC CACCAACTGG      350
   TCTGAGTCCC GTTTCGGTGA AATCATCAAG GAAGTCACCA ACTTCATCAA      400
   GAAGGTCGGC TACGACCCCA AGGGTGTCCC ATTCGTCCCA ATCTCTGGCT      450
   TCAACGGTGA CAACATGATT GAGCCCTCCA CCAACTGCCC ATGGTACAAG      500
25 GGATGGAACA AGGAGACCAA GGCCGGTGCC AAATCCTCTG GTAAGACCCT      550
   CCTTGAGGCC ATCGATGCCA TTGACATGCC CACTCGTCCC ACCGACAAGC      600
   CTCTCCGTCT CCCACTCCAG GATGTCTACA AGATCTCTGG TATCGGAACA      650
   GTACCAAGTC GTCGTGTTGA GACTGGTATC ATCAAGCCTG GTATGGTTGT      700
   CACTTTCGCC CCCGCCAACG TCACCACTGA AGTCAAGTCC GTCGAAATGC      750
30 ACCACCAGCA GCTCGTTCAG GGTGTTCCCG GTGACAACGT TGGCTTCAAC      800
   GTCAAGAACG TCTCTGTCAA GGAAGTCCGC CGTGGTAACG TTGCCGGTGA      850
   TTCCAAGAAC GACCCACCAT CTGGCTGCGC CTCTTTCAAG GCCCAGGTCA      900
   TCGTCCTCAA CCACCCCGGC CAGATCGGTG CTGGTTACGC CCCAGTCCTC      950
   GACTGCCACA CTGCCCACAT TGCTTGCAAG TTCTCTGAGC TTCTTGAGAA      1000
35 GATTGACCGC CGTACTGGTA AATCCGTCGA AACCAGCCCT AAGTTCGTCA      1050
   AGTCTGGTGA TGCCGCTATT GCCACCATGG TTCCATCCAA GCCCATGTGC      1100
   GTTGAGGCTT TCACTGACTA CCCACCACTT GGTGTTTCG CCGTCCGTGA      1150
C

```

## 2) INFORMATION FOR SEQ ID NO: 784

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 979 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mucor circinelloides*  
 (B) STRAIN: ATCC 38592

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 784

```

CCAAGTGGTC TCAAGATCGT TACAACGAAA TTGTCAAGGA AGTTTCCGGT      50
TTCATCAAGA AGATCGGTTT CAACCCCAAG TCCGTTCCCT TCGTTCCCAT      100
60 TTCTGGCTGG CACGGTGATA ACATGTTGGA TGAATCCACC AACATGCCCT      150

```

	GGTTCAAGGG	ATGGAACAAG	GAGACCAAGG	CCGGTTCCAA	GA CTGGTAAG	200
	ACTCTCCTCG	AAGCCATCGA	TGCCATTGAG	CCCCCTGTCC	GTCCTTCTGA	250
	CAAGCCTCTC	CGTCTTCCTC	TTCAAGATGT	CTACAAGATT	GGTGGTATTG	300
	GTACAGTTCC	CGTCGGTCGT	GTTGAAACTG	GTACTATCAA	GGCTGGTATG	350
5	GTTGTCAACT	TCGCTCCCGC	TGCTGTCACC	ACTGAAGTTA	AGTCCGTTGA	400
	AATGCATCAC	GAAACCCTCT	CTGAAGGTCT	CCCCGGTGAC	AACGTTGGTT	450
	TCAACGTCAA	GAACGTCTCC	GTCAAGGATA	TCCGTCGTGG	TAACGTCTGT	500
	TCCGACTCCA	AGAACGATCC	CGCTAAGGAA	TCTGCCTCTT	TCACTGCTCA	550
	AGTTATTATC	TTGAACCATC	CCGGTCAAAT	CTCTGCTGGT	TACGCACCAG	600
10	TTCTCGATTG	TCACACTGCT	CACATCGCCT	GTAAGTTCTC	TGAACTCATT	650
	GAGAAGATTG	ATCGTCGTTT	CGGTAAGTAC	CTGCATCTGT	CAGAATTGAA	700
	GGTCCGCCGT	TATAGCAAAG	GCTGGGTTTA	AATGTTGGGG	TTTGTCTGAT	750
	CTATAATGAT	GATTGCTCCT	TCAATTTTGT	ACATAATTTG	ATGATCTGAA	800
	TTGTGTTGCT	AACGTCGCAT	TTGCTTCTTT	TGCTTCTTTT	GCATGTAGGT	850
15	AAGAAGATTG	AGATAGCTCC	CAAAGTAAGT	ATTACGATTG	ATGGACAATT	900
	AAAATAGAAT	ACTAACAATT	ATTGTTTATA	GTTCTGCAAG	TCTGGTGACT	950
	CTGCTATCGT	CAAGATGGTT	CCCTCCAAG			979

20

## 2) INFORMATION FOR SEQ ID NO: 785

## (i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 1099 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

30

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Phialaphora verrucosa*
- (B) STRAIN: ATCC 38561

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 785

	GGACAAGCTG	AAGGCCGAAC	GTGAGCGTGG	TATCACCATC	GATATCGCGC	50
	TCTGGAAGTT	CGAGACTCCC	AAATACTTCG	TCACCGTCAT	TGATGCCCCCT	100
	GGTCATCGTG	ACTTCATCAA	GAACATGATC	ACTGGTACCT	CCCAGGCTGA	150
40	TTGTGCCATC	CTCATCATTG	CCGCCGGTAC	CGGTGAGTTC	GAAGCCGGTA	200
	TCTCCAAGGA	TGGCCAGACC	CGTGAGCACG	CTCTCCTCGC	CTACACCCTA	250
	GGTGTGAAGC	AGCTTATCGT	CGCCATCAAC	AAGATGGACA	CCGCCAAATG	300
	GTCCGAGGAT	CGGTTCAACG	AAATCATCAA	GGAGACTTCC	AACTTCATCA	350
	AGAAGGTCGG	ATACAACCCC	AAGTCCGTCC	CGTTCGTGCC	CATCTCCGGT	400
45	TTCAACGGTG	ACAACATGAT	CGACGTCTCC	TCCAACGCCC	CCTGGTACAA	450
	GGGTTGGGAG	AAGGAGACCA	AGGCCGGCAA	GGCCACTGGC	AAGACCCTCC	500
	TCGAGGCCAT	CGACGCGATT	GACCCTCCTA	CTCGTCCCAC	CGACAAGCCC	550
	CTCCGTCTCC	CTCTCCAGGA	TGTCTACAAG	ATCTCTGGTA	TCGGCACGGT	600
	GCCCGTTGGT	CGTGTTGAGA	CCGGTACCAT	CAAAGCCGGT	ATGGTCGTCA	650
50	CCTTCGCTCC	CGCCAACGTC	ACCACTGAAG	TCAAGTCCGT	CGAAATGCAC	700
	CACGAACAGC	TCGCCGAAGG	TGTTCCAGGT	GACAATGTCT	GCTTCAACGT	750
	CAAGAACGTC	TCCGTCAAGG	AGGTTCTGTC	TGGAAACGTT	GCCGGTGACT	800
	CCAAGAATGA	CCCCCCCCAAG	GGTGCCGACT	CCTTCAACGC	TCAGGTCATC	850
	GTCCTCAACC	ACCCTGGTCA	GGTCGGTGCC	GGCTACGCCC	CGGTCTTGGA	900
55	TTGCCACACT	GCCCACATTG	CTTGCAAGTT	CTCTGAGCTC	CTCGAGAAGA	950
	TCGATCGTCG	TACCGGCAAG	TCCATGGAAA	ACAACCCCAA	GTTTATCAAG	1000
	TCTGGTGATG	CTGCCATCGT	GAAGATGGTT	CCCAGCAAGC	CTATGTGCGT	1050
	TGAGGCCTTC	ACCGACTATC	CTCCTCTTGG	TCGTTTCGCC	GTCCGTGAC	1099

60

## 2) INFORMATION FOR SEQ ID NO: 786

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 750 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Saksenaea vasiformis*  
 (B) STRAIN: ATCC 60625

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 786

	ACCACCAAGT	GGTCTGAGGC	CCGTTACCAG	GAGATCATCA	AGGAGACCTC	50
	CGGYTTCATC	AAGAAGGTCG	GCTTCAACCC	CAAGCACGTT	CCCTTCGTGC	100
20	CCATCTCCGG	TTTCAACGGT	GACAACATGA	TCGACGTCTC	CACCAACTGC	150
	CCCTGGTACA	AGGGTTGGGA	GAAGGAGACC	AAGGCCAAGG	CCACCGGCAA	200
	GACCMTCCTC	GAGGCCATTG	ACGCCATCGA	CCCCCYAGY	CGTCCCACCG	250
	ACAAGCCCCT	YCGTCTTCCC	CTMCAGGATG	TYTACAAGAT	TGGCGGTATT	300
	GGCAGCGTTC	CCGTCGGTCC	TGTYGAGACC	GGTRCCATCA	AGGGTGGCAT	350
25	GGTCGTCACC	TTCCCCCCCC	CTGGTGTCAC	CACTGAGGTC	AAGTCCGTCG	400
	AGATGCACCA	CGAGCAGCTC	GCCGAGGGTS	TCCCCGGTGA	CAACGTCGGC	450
	TTCAACGTCA	AGAACGTCTC	CGTCAAGGAG	ATCCGTCGTG	GCAACGTTGC	500
	CGGTGACTCC	AAGAACGACC	CCCCCAAGGG	CTGCGACAGC	TTCAACGCCC	550
	AGGTCATCGT	CCTCAACCAC	CCCGGTCAGG	TCGGCGCCGG	CTACGCGCCS	600
30	GTCCCTSGACT	GCCACACTGC	TCACATTGCC	TGCAAGTTCT	CTGAGCTCCT	650
	CGAGAAGATC	GACCGCCGTT	CCGGCAAAGT	CATCGAGTCC	GGCCCCAAGT	700
	TCATCAAGTC	TGGTGACGCC	GCCATCGTCA	AGATGGTTCC	CTCCAAGCCC	750

## 2) INFORMATION FOR SEQ ID NO: 787

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1084 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Syncephalastrum racemosum*  
 (B) STRAIN: ATCC 32330

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 787

	GCTCAAGGCC	GAGCGTGAGC	GTGGTATCAC	CATCGATATC	GCTCTCTGGA	50
	AGTTCGAGAC	CCCCAAGTAC	CACGTCACCG	TCATTGATGC	CCCCGGCCAT	100
	CGTGATTTCA	TCAAGAACAT	GATCACTGGT	ACCTCGCAGG	CTGACTGCGG	150
55	TATCCTCATC	ATTGCCGCCG	GTA CTGGTGA	GTTCGAGGCT	GGTATCTCCA	200
	AGGATGGCCA	GACCCGTGAG	CACGCTCTGC	TTGCCTTCAC	CCTCGGTGTC	250
	CGTCAGCTGA	TCGTCCGCCAT	CAACAAGATG	GACTCGACCA	AGTACTCTGA	300
	GGCCCGTTAC	AACGAAATCG	TCAAGGAGGT	CTCCACCTTC	ATCAAGAAGA	350
	TCGGTTTCAA	CCCCAAGTCC	GTTCCCTTCG	TCCCCATCTC	TGGCTGGAAC	400
60	GGTGACAACA	TGTTGGAGGA	GTCTCCAAC	ATGCCCTGGT	TCAAGGGCTG	450

```

GAAGAAGGAG ACCAAGGCTG GCGAGAAGTC CGGCAAGACC CTCCTTGAGG 500
CCATTGACAA CATTGACCCC CCGGTCCGTC CCTCGGACAA GCCCCTCCGT 550
CTTCCCCTCC AGGATGTCTA CAAGATCGGT GGTATCGGCA CAGTCCCCGT 600
CGGTCTGTGTC GAGACTGGTG TCATCAAGGC TGGTATGGTC GTGACCTTCG 650
5 CCCCCGCCAA CGTCACCACT GAAGTCAAGT CCGTCGAGAT GCACCACGAG 700
CAGCTCGTCG AGGGTGTCCC CGGTGACAAC GTCGGTTTCA ACGTCAAGAA 750
CGTTTCCGTC AAGGATATCC GCCGTGGTAA CGTCTGCTCT GACTCCAAGA 800
ACGACCCCGC CAAGGAGTCT GCCTCGTTCA CCGCCAGGT CATCGTCCTG 850
AACCACCCCG GTCAGATCGG TGCCGGTTAC GCCCCGGTTC TTGACTGCCA 900
10 CACCGCTCAC ATTGCCTGCA AGTTCGCTGA GCTCCTCGAG AAGATCGACC 950
GTCGTTCCGG YAAGAAGCTC GAAGAGTCCC CCAAGTTCGT CAAGTCGGGT 1000
GACTCCGCCA TCGTCAAGAT GGTTCCTCC AAGCCCATGT GCGTTGAGGC 1050
CTACACTGAG TACCCCCCTC TTGGCCGTTT CGCC 1084

```

15

## 2) INFORMATION FOR SEQ ID NO: 788

```

20 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1155 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

25 (ii) MOLECULE TYPE: Genomic DNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Trichophyton tonsurans
        (B) STRAIN: ATCC 56185

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 788

```

```

GCTCAAGGCC GAGCGTGAGC GTGGTATCAC CATCGATATC GCCCTCTGGA 50
AGTTCGAGAC CCCCAAGTAC AATGTCACCG TCATTGGTAT GTTTTCTTT 100
35 ACCTTTCCCC TCCATCGTCT TGCTGTGCCA TAACTAACGA GAGTAGACGC 150
CCCCGGTCCAC CGTGACTTCA TCAAGAACAT GATCACTGGT ACCTCCCAGG 200
CTGACTGTGC TATTCTCATC ATTGCTGCCG GTACTGGTGA GTTCGAGGCT 250
GGTATCTCCA AGGATGGCCA GACCCGTGAG CACGCTCTGC TCGCCTTCAC 300
CCTTGGTGTC AAGCAGCTCA TCGTTGCCAT CAACAAGATG GACACCACCA 350
40 ACTGGTCCGA GGACCGTTTC AAGGAAATCA TCAAGGAAGT CACCAACTTC 400
ATCAAGAAGG TTGGCTACGA CCCCAGGGT GTTCCATTCG TTCCAATCTC 450
TGGTTTCAAC GGTGACAACA TGATTGAGGC CTCCAGCAAC TGCCCATGGT 500
ACAAGGGATG GAACAAGGAG ACCAAGGCCG GTGGTGCCAA GACTGGCAAG 550
ACCCTYCTCG AGGCCATCGA TGCCATCGAC ATGCCAACCC GTCTTACCGA 600
45 CAAGCCCCTY CGTCTCCAC TCCAGGATGT CTACAAGATC TCTGGTATCG 650
GAAGTGTACC AGTCGGTCGT GTTGAGACCG GTATCATCAA GCCTGGTATG 700
GTCGTCACCT TCGCCCCGTC CAACGTCACC ACTGAAGTCA AGTCCGTCGA 750
AATGCACCAC CAGCAGCTTC AGCAGGGTGT CCCCAGGTGAC AACGTCGGCT 800
TCAACGTCAA GAACGTTTCC GTCAAGGAAG TCCGCCGTGG TAACGTTGCC 850
50 GGTGACTCCA AGAACGACCC ACCATCCGGC TGTGCCTCCT TCAACGCCCA 900
GGTCATCGTC CTCAACCACC CCGGCCAGAT CGGTGCTGGT TACGCCCCAG 950
TCCTCGACTG CCACACTGCT CACATTGCTT GCAAGTTCGC TGAGCTCCTC 1000
GAGAAGATTG ACCGCCGTAC CGGTAAATCC GTCGAAGCCA ACCCAAGTT 1050
CGTCAAGTCT GGTGATGCCG CTATCGCCAA GATGGTTCCC TCCAAGCCCA 1100
55 TGTGCGTTGA GGCTTTCACT GACTACCCCC CACTTGGTCG TTTCCGCCGTC 1150
CGTGA 1155

```

60 2) INFORMATION FOR SEQ ID NO: 789

653

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1138 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichophyton mentagrophytes*  
 (B) STRAIN: ATCC 8125

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 789

```

15 TCAAGGCCGA GCGTGAGCGT GGTATCACCA TCGATATCGC CCTCTGGAAG      50
   TTCGAGACCC CCAAGTACAA TGTCACCGTC ATTGGTATGT TTCTCTTTAC      100
   CTTTCCCCTC CATCGTCTTG CTGTGCCATA ACTAACGAGA GTAGACGCCC      150
   CCGGTCACCG TGACTTCATC AAGAACATGA TCACTGGTAC CTCCCAGGCT      200
20 GACTGTGCTA TTCTCATCAT TGCTGCCGGT ACTGGTGAGT TCGAGGCTGG      250
   TATCTCCAAG GATGGCCAGA CCCGTGAGCA CGCTCTGCTC GCCTTCACCC      300
   TTGGTGTCAA GCAGCTCATC GTTGCCATCA ACAAGATGGA CACCACCAAC      350
   TGGTCCGAGG ACCGTTTCAA GGAAATCATC AAGGAAGTCA CCAACTTCAT      400
   CAAGAAGGTT GGCTACGACC CCAAGGGTGT TCCATTCTGT CCAATCTCTG      450
25 GTTTCACCGG TGACAACATG ATTGAGGCCT CCACCAACTG CCCATGGTAC      500
   AAGGGATGGA ACAAGGAGAC CAAGGCCGGT GGTGCCAAGA CTGGCAAGAC      550
   CCTCCTCGAG GCCATCGATG CCATCGACAT GCCAACCCGT CCTACCGACA      600
   AGCCCCTCCG TCTCCCACTC CAGGATGTCT ACAAGATCTC TGGTATCGGA      650
   ACTGTACCAG TCGGTCGTGT TGAGACCGGT ATCATCAAGC CCGGTATGGT      700
30 CGTCACCTTC GCCCCTGCCA ACGTCACCAC TGAAGTCAAG TCCGTCGAAA      750
   TGCACCACCA GCAGCTTCAG CAGGGTGTCC CCGGTGACAA CGTCGGCTTC      800
   AACGTCAAGA ACGTTTCCGT CAAGGAAGTC CGCCGTGGTA ACGTTGCCGG      850
   TGACTCCAAG AACGACCCAC CATCCGGCTG TGCTCCTTC AACGCCCAGG      900
   TCATCGTCTT CAACCACCCC GGCCAGATCG GTGCTGGTTA CGCCCCAGTC      950
35 CTCGACTGCC AACTGTCTCA CATTGCTTGC AAGTTCGCTG AGCTCCTCGA      1000
   GAAGATTGAC CGCCGTACCG GTAAATCCGT CGAAGCCAAC CCCAAGTTCTG      1050
   TCAAGTCTGG TGATGCCGCT ATCGCCAAGA TGGTTCCATC CAAGCCTATG      1100
   TGCGTTGAGG CTTTCACTGA CTACCCCCCA CTTGGTCTG      1138

```

40

## 2) INFORMATION FOR SEQ ID NO: 790

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 748 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bipolaris hawaiiensis*  
 (B) STRAIN: ATCC 26067

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 790

```

60 CACCACCAAG TGGTCTGAGG AGCGTTACCA GGAAATCATC AAGGAGACCT      50
   CCAACTTCAT CAAGAAGGTC GGCTACAACC CCAAGCACGT TCCCTTCGTC      100
   CCCATCTCTG GTTTCACCGG AGACAACATG ATTGAGGCTT CCACCAACTG      150

```

	CCCCTGGTAC	AAGGGTTGGG	AGAAGGAGAC	CAAGGCCAAG	GCCACTGGTA	200
	AGACCCTTCT	CGAGGCCATC	GACGCCATCG	ACCCCCCTGT	CCGTCCTACC	250
	GACAAGCCCC	TCCGCCTTCC	CCTCCAGGAT	GTGTACAAGA	TTGGTGGTAT	300
	TGGCACGGTT	CCCGTCGGTC	GTGTCGAGAC	CGGTATCATC	AAGCCCGGTA	350
5	TGGTCGTCAC	CTTCGCCCCC	GCTGGTGTCA	CCACTGAAGT	CAAGTCCGTC	400
	GAGATGCACC	ACGAGCAGCT	TGCCGAGGGT	GTCCCCGGTG	ACAACGTCGG	450
	CTTCAACGTC	AAGAACGTCT	CCGTCAAGGA	GATCCGTCGT	GGTAACGTTG	500
	CCGGTGACTC	CAAGAACGAC	CCCCCAAGG	GTTGCGAGTC	CTTCAACGCC	550
	CAGGTCATCG	TCCTCAACCA	CCCCGGTCAG	GTCGGTGCCG	GTTACGCACC	600
10	AGTCCTTGAC	TGCCACACTG	CCCACATTGC	TTGCAAGTTC	TCCGAGCTCC	650
	TCGAGAAGAT	TGACCGCCGT	ACCGGAAAGT	CTGTTGAGAA	CTCCCCCAAG	700
	TTCATCAAGT	CCGGTGACGC	CGCCATCGTC	AAGATGGTTC	CCTCCAAG	748

15

## 2) INFORMATION FOR SEQ ID NO: 791

## (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 958 bases
20	(B) TYPE: Nucleic acid
	(C) STRANDEDNESS: Double
	(D) TOPOLOGY: Linear

25

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

	(A) ORGANISM: <i>Aspergillus fumigatus</i>
	(B) STRAIN: ATCC 14110

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 791

	CGCTATTGTC	GTTGTTGCTG	CCTCCGACGG	TCAGATGTAG	GTGGAACATC	50
	TTGGGAAATA	CGTCGTAAAA	CACGGCGCTT	ACGTTTTCGC	GAATAGGCCC	100
	CAGACTCGTG	AGCATTGCT	GCTCGCCCGC	CAGGTTGGTG	TCCAGAAGAT	150
35	CGTTGTCTTC	GTCAACAAAA	TCGATGCTAT	TGATGATCCG	GAGATGCTGG	200
	AACTGGTCGA	ACTCGAGATG	CGTGAGCTGC	TGAACAGCTA	CGGTTTCGAG	250
	GGTGAAGAGA	CTCCGATCAT	TTTCGGTTCC	GCTCTCTGTG	CTCTCGAAGG	300
	ACGCCGTGAC	GACATCGGTA	AAGACAGAAT	TGAGCAGCTT	ATGAACGCTG	350
	TCGACACCTG	GATCCCCACT	CCTCAGCGTG	ACCTCGACAA	ACCTTTCTTG	400
40	ATGTCTGTCTG	AGGAAGTGTT	CTCTATCGCC	GGCCGTGGTA	CCGTGGCTTC	450
	TGGTCGTGTC	GAGCGTGGTA	TCTTGAAGAA	GGACTCTGAG	GTTGAGATTG	500
	TTGGAGGCTC	CTTCGAACCC	AAGAAGACCA	AAGTCACCGA	CATTGAAACC	550
	TTCAAGAAGA	GCTGTGATGA	ATCGCGTGCT	GGTGACAAC	CTGGTCTCCT	600
	CCTGCGTGGT	ATCCGACGTG	AAGACGTCAA	GCCTGGTATG	GTCATTGCTG	650
45	TTCCCGGCAG	CACCAAAGCT	CACGACAAGT	TCCTCGTCTC	CATGTACGTC	700
	CTGACCGAGG	CGGAGGGTGG	TCGTCTGACT	GGCTTCGGTG	CCAACTACCG	750
	TCCCCAAGTC	TTCATCCGTA	CTGCAGGTAA	GTTCCCGCAC	ACCGTGTCCA	800
	GATCTTCCGA	GAGATTAGCG	ATATATGCTA	ATGATTCATC	AGACGAGGCT	850
	GCTGACCTCA	GCTTCCCTGA	CGGCGACCAA	TCTCGCAGAG	TTATGCCTGG	900
50	TGACAACGTC	GAGATGATCC	TGAAGACCCA	CCACCCTGTT	GCTGCTGAGG	950
	CTGGTCAA					958

55

## 2) INFORMATION FOR SEQ ID NO: 792

## (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 936 bases
60	(B) TYPE: Nucleic acid
	(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trichophyton mentagrophytes*  
(B) STRAIN: ATCC 8125

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 792

```

10 CGTTGTCGCA GCTTCTGACG GTCAAATGTA ATTGAATGCC CGCCCAGACG      50
   GATGAAAGGA TTTGACGTTT CTAACATCAT TCTAGGCCTC AGACCAGAGA      100
   ACATTTGCTC CTTGCCCCGCC AGGTCGGTGT CCAGAAGCTG GTCGTTTTTCG      150
   TTAACAAGGT CGATGCCGTT GAGGACCCAG AGATGTTGGA GCTTGTCGAA      200
15 CTTGAAATGC GTGAACTCCT CAGCCACTAC GGTTTCGAGG GTGAGGAGAC      250
   CCCCATCATT TTTGGCTCTG CTCTCTGTGC CCTCGAGTCC CGTCGACCTG      300
   AGCTTGGTGT CGAGAAGATT GACGAGCTAT TGAACGCCGT CGACACCTGG      350
   ATCCCCACCC CCGAGCGCGC CACTGATAAG CCTTTCCTCA TGTCCATTGA      400
   GGAAGTGTTT TCTATCTCTG GTCGTGGTAC CGTCGTCTCC GGTCTGTGTTG      450
20 AGCGTGATAT CCTCAAGAAG GATTCCGACG TCGAAATTGT TGGTGGCTCT      500
   ACCACCCCTA TCAAGACCAA GGTCACAGAT ATCGAAACCT TCAAGAAGTC      550
   CTGCGATGAA TCTCGAGCTG GTGACAACCTC TGGTCTCCTT CTCCGAGGTA      600
   TCAAGCGTGA GGACTTGAAG CGTGGAATGG TTGTTGCTGC CCCC GGATCC      650
   ACCAAGGCTC ACACCGACTT CATGGTCTCC CTCTACGTCC TGACTGAGGC      700
25 TGAGGGTGGT CGTTCCAACG GCTTCACCCA CAAGTACCGC CCCC AAATGT      750
   TCATCCGTAC TGCTGGTATG TAACCCAAGT TTCCGCTATT TACTAAGTAG      800
   ATCATTGCTA ACTTGTATTT CCTTCCGTAG ACGAAGCCGC ATCTTTCAGC      850
   TGGCCTGGAG AAGACCAAGA CAAGAAGGCT ATGCCTGGTG ACAACGTCGA      900
   GATGATTTGC AAGACCCTCC ACCCCATTGC TGCCGA      936
30

```

2) INFORMATION FOR SEQ ID NO: 793

35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

40 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 793

45 TTATTGTTGC TGCTGGTACT

20

2) INFORMATION FOR SEQ ID NO: 794

50 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

55 (ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

- 60 (A) ORGANISM: *Trypanosoma cruzi*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 794

GACGACAAGT CGGTGAACTT

20

2) INFORMATION FOR SEQ ID NO: 795

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma cruzi*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 795

ACTTGCACGC GATGTGGCAG

20

2) INFORMATION FOR SEQ ID NO: 796

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 796

GGTCCAATGC CWCAAACWAG A

21

2) INFORMATION FOR SEQ ID NO: 797

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 797

CATTAAGAAT GGYTTATCTG TSKCTCT

27

2) INFORMATION FOR SEQ ID NO: 798



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 798

TGGTTGTCCC AGCCGATCGT TT

22

## 2) INFORMATION FOR SEQ ID NO: 799

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 799

ACCTGTGAAT ACAAGCAATC T

21

## 2) INFORMATION FOR SEQ ID NO: 800

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 800

GATGAAATCT TCAACGAAGT TGAT

24

## 2) INFORMATION FOR SEQ ID NO: 801

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid

658

(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 801

ACAACACCGA GAAGATCCCA

20

2) INFORMATION FOR SEQ ID NO: 802

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 802

TTGCCATTTC TGGTTTCGTT

20

2) INFORMATION FOR SEQ ID NO: 803

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 803

ACTTCAGTGG TAACACCAGC

20

2) INFORMATION FOR SEQ ID NO: 804

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 804

CCTGGGACGG CCTCTGGCAT

20

10

2) INFORMATION FOR SEQ ID NO: 805

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

15 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 805

25

CTCTTGTCCTCA TCTTAGCAGT

20

30 2) INFORMATION FOR SEQ ID NO: 806

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 bases

(B) TYPE: Nucleic acid

35 (C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

40 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 806

45 AGCATCACCA GACTTGATAA G

21

2) INFORMATION FOR SEQ ID NO: 807

50

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

55 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

60 (A) ORGANISM: *Cryptosporidium parvum*

660

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 807

AAAGTGGCTT CAAAGGTTGC

20

5

2) INFORMATION FOR SEQ ID NO: 808

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 808

GCITTAIWRG CATTAGAARA YCCA

24

20

2) INFORMATION FOR SEQ ID NO: 809

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 809

TCTTCCTGTW GCAACTGTTC CTCT

24

35

2) INFORMATION FOR SEQ ID NO: 810

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 810

AGAGMWACAG ATAARSCATT CTTA

24

50

2) INFORMATION FOR SEQ ID NO: 811

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

661

60

55

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 811

5

TRAARTAGAA TTGTGGTCTR TATCC

25

10 2) INFORMATION FOR SEQ ID NO: 812

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

(B) TYPE: Nucleic acid

15

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 812

GTIACIGGIT CYTYRARRTT ICCICC

26

25

2) INFORMATION FOR SEQ ID NO: 813

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 bases

30

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 813

AATCYGTYGA AATGCAYCAC GA

22

40

2) INFORMATION FOR SEQ ID NO: 814

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

45

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 814

GCIGGCACGT ACACIGCCTG

20

55

2) INFORMATION FOR SEQ ID NO: 815

60 (i) SEQUENCE CHARACTERISTICS:

662

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 815

10 TGGTGCATYT CKACRGACTT

20

2) INFORMATION FOR SEQ ID NO: 816

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 816

25

GCTACGACGA GATCAAGGGC

20

30 2) INFORMATION FOR SEQ ID NO: 817

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 817

40

TGGAAGAAGG CCGAGGAGTT

20

45

2) INFORMATION FOR SEQ ID NO: 818

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 818

55

AGCCGGGCTG GATCTTCTTC

20

60

## 2) INFORMATION FOR SEQ ID NO: 819

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 819

TCGAGCTTCT GGAGGAAGAG

20

## 2) INFORMATION FOR SEQ ID NO: 820

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma brucei*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 820

GAAGGAGGTG TCTGCTTACA C

21

## 2) INFORMATION FOR SEQ ID NO: 821

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma brucei*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 821

GGCGCAAACG TCACCACATC A

21

## 2) INFORMATION FOR SEQ ID NO: 822

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid

664

(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma brucei*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 822

CGGCGGATGT CCTTAACAGA A

21

2) INFORMATION FOR SEQ ID NO: 823

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 823

GAGCGGTATG AYGAGATTGT

20

2) INFORMATION FOR SEQ ID NO: 824

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 824

GGCTTCTGCG GCACCATGCG

20

2) INFORMATION FOR SEQ ID NO: 825

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 825

ATGAGCARCG SAACCATCGT TCAGTG

26



## 2) INFORMATION FOR SEQ ID NO: 826

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 826

15 TCGATCGTGC CGACCATGTA GAACGC

26

## 2) INFORMATION FOR SEQ ID NO: 827

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 446 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 25 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:  
 30 (A) ORGANISM: *Clostridium novyi*  
 (B) STRAIN: ATCC 19402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 827

35 CACCAACTTG CTAAATGGGG AGATGCCCGAG ATTGTTGTAT ATATAGGCTG 50  
 TGGAGAACGT GGAAATGAAA TGACAGATGT TCTTAATGAG TTTCCAGAAC 100  
 TTAAAGATCC TAAGACTGGC AAATCAATAA TGGAAAGAAC AGTTTAAATA 150  
 GCAAATACTT CTAATATGCC AGTTGCAGCC CGTGAAGCTT GTATATATAC 200  
 AGGAATCACA ATAGCAGAAT ATTTTAGAGA TATGGGATAT TCAGTAGCAC 250  
 40 TTATGGCGGA TTCCACTTCA CGTTGGGCAG AGGCATTAAG AGAAATGTCT 300  
 GGAAGACTTG AAGAAATGCC TGGTGATGAA GGTACCCAG CTTATTTAGG 350  
 ATCAAGACTT GCTGATTTCT ATGAAAGAGC TGGAAAAGTT GTGTGTTTAG 400  
 GAGACGATGA AAGAGAAGGT GCCATTACTG CAATAGGTGC TGTATC 446

45

## 2) INFORMATION FOR SEQ ID NO: 828

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 445 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

55 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Clostridium difficile*  
 (B) STRAIN: 9689

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 828

```

5  CAGCATCAGC TTGCTAAATG GGCAGATGCA GATATAGTTG TATATATAGG 50
   CTGTGGCGAG CGTGGAAATG AAATGACAGA TGTTCTTCTT GAATTTCCCTG 100
   AATTAAAAGA CCCAAGAACA GGCGAGTCAC TTATGCAAAG AACTGTGCTT 150
   ATAGCAAATA CATCAGATAT GCCGGTTGCT GCACGTGAAG CTTCTATATA 200
   CACTGGTATT ACAATAGCTG AATATTTTAG AGATATGGGA TATAGTGTG 250
   CACTTATGGC AGACTCTACA TCAAGATGGG CTGAGGCTCT TAGAGAGATG 300
   AGTGGTCGTT TAGAGGAGAT GCCTGGTGAA GAAGGTTATC CTGCATACTT 350
10 AGGTTACAGT CTTGCTCAAT TCTATGAGAG AGCAGGAAAG GTAAATTGTC 400
   TAGGTATGGA TGAAAGAGAA GGAACACTTA CAGCAATTGG TGCAG 445

```

## 15 2) INFORMATION FOR SEQ ID NO: 829

## (i) SEQUENCE CHARACTERISTICS:

```

20 (A) LENGTH: 445 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
   (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 25 (vi) ORIGINAL SOURCE:

```

   (A) ORGANISM: Clostridium septicum
   (B) STRAIN: ATCC 12464

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 829

```

30 ATGCTATAGC TAAATGGGGA GACAGCGAAA TAGTTGTTTA CGTTGGATGT 50
   GGAGAACGTG GTAACGAAAT GACAGACGTT CTTAACGAAT TCCCAGAACT 100
   TATTGACCCA AAAACTGGGG AAAGTTTAAAT GAAGAGAACA GTACTTATAG 150
   CTAATACTTC AAACATGCCA GTTGCTGCTA GAGAAGCTTG CATATACACA 200
35 GGTATTACAA TAGCTGAATA CTTCAGAGAT ATGGGATACT CAGTATCTAT 250
   AATGGCTGAT TCAACTTCAA GATGGGCAGA AGCATTAAGA GAAATGTCAG 300
   GTAGACTTGA AGAAATGCCA GGTGATGAAG GATATCCAGC GTACTTAGGA 350
   TCAAGACTTG CTGATTATTA CGAAAGAGCA GGTAAGGTTG TTTGTCTAGG 400
40 TAAAGATGGT AGAGAAGGTG CTGTAACAGC AATTGGAGCT GTATC 445

```

## 2) INFORMATION FOR SEQ ID NO: 830

## 45 (i) SEQUENCE CHARACTERISTICS:

```

   (A) LENGTH: 444 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
50 (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

```

55 (A) ORGANISM: Clostridium botulinum
   (B) STRAIN: 20:3.1

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 830

```

60 TCAAATTGCT AAATGGGGAG ATGCAGAAAT CGTTGTTTAC GTTGGATGCG 50
   GAGAACGTGG TAACGAAATG ACAGACGTTG TTAATGAGTT CCCAGAACTT 100

```

	ATTGACCCTA	AGACTGGCGA	AAGCTTAATG	AAGAGAACAG	TTCTTATAGC	150
	TAATACTTCA	AACATGCCAG	TTGCAGCGAG	AGAAGCTTCA	ATATATACAG	200
	GTATCACAAT	AGCTGAATAT	TTCAGAGATA	TGGGATATGC	AGTATCAATA	250
	ATGGCTGACT	CAACTTCAAG	ATGGGCTGAG	GCATTAAGAG	AAATGTCTGG	300
5	TAGACTTGAA	GAAATGCCTG	GTGATGAAGG	ATATCCAGCT	TACCTTGGAT	350
	CAAGACTTGC	TGATTACTAT	GAAAGAGCTG	GTAAGGTTGA	ATGTTTAGGT	400
	AATGATGGAA	GAATTGGTTC	TATAACAGCA	ATCGGTGCGG	TATC	444

## 10 2) INFORMATION FOR SEQ ID NO: 831

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 456 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Clostridium perfringens*  
(B) STRAIN: ATCC 13124

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 831

25	TAGTTCAGCA	CCAAGTTGCT	AAATGGGGAG	ATACTGAGAT	AGTTGTTTAC	50
	GTTGGATGTG	GAGAACGTGG	TAACGAGATG	ACAGACGTTT	TTAACGAATT	100
	CCCAGAACTT	AAAGACCCTA	AACTGGGGGA	AAGCTTAATG	AAGAGAACAG	150
	TTCTTATTGC	TAATACATCT	AACATGCCAG	TTGCTGCCAG	AGAAGCATCA	200
30	ATATATACTG	GTATAACAAT	AGCAGAGTAT	TTCAGAGATA	TGGGATACTC	250
	AGTATCAATC	ATGGCTGACT	CAACTTCACG	TTGGGCAGAG	GCTTTAAGAG	300
	AAATGTCAGG	AAGACTTGAA	GAAATGCCAG	GAGACGAAGG	TTACCCAGCA	350
	TACTTAGGAT	CAAGACTTGC	TGATTACTAT	GAAAGAGCTG	GTAAAGTTGT	400
	AGCTTTAGGT	AAAGATGGAA	GAGAAGGAGC	TGTTACAGCT	ATCGGAGCAG	450
35	TATCCC					456

## 2) INFORMATION FOR SEQ ID NO: 832

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 444 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Clostridium tetani*  
(B) STRAIN: ATCC 19406

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 832

	CCAAGTTGCA	AAATGGGCTG	ATGCTCAAAT	AGTTGTGTAC	ATAGGATGTG	50
55	GAGAACGTGG	AAATGAAATG	ACAGACGTTT	TAAATGAGTT	CCCAGAATTA	100
	AAGGATCCTA	AAACCGGGGA	ATCTTTAATG	AAAAGAACTG	TGTTAATAGC	150
	AAATACATCT	AATATGCCTG	TTGCAGCTAG	AGAAGCATCT	ATATATACTG	200
	GTATAACAAT	AGGGGAATAT	TTTAGAGATA	TGGGATATTC	AATAGCACTA	250
	ATGGCAGATT	CGACTTCTAG	ATGGGCAGAG	GCTCTAAGAG	AAATGTCTGG	300
60	AAGACTAGAG	GAGATGCCAG	GTGAAGAAGG	TTATCCAGCT	TATTTAGGAT	350

CTAGATTAGC	AGAGTTCTAT	GAAAGAGCAG	GTAATGTTAT	ATGTTTAGGT	400
CAGGATGGAA	GAGAAGGAGC	ATTAACAGCT	ATAGGAGCAG	TTTC	444

5

## 2) INFORMATION FOR SEQ ID NO: 833

## (i) SEQUENCE CHARACTERISTICS:

10

- (A) LENGTH: 1786 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

15

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pyogenes*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 833

20

TGAACCAAGG	AAAAATAATA	ACCGTTTCGG	GACCTCTTGT	TGTGGCTTCT	50
GGGATGCAAG	AAGCTAATAT	TCAAGATATT	TGTCGTGTGG	GACATCTTGG	100
CTTAGTCGGA	GAAATTATTG	AAATGCGTCG	CGATCAAGCG	TCTATTCAGG	150
TTTATGAGGA	AACATCAGGG	ATCGGTCCAG	GAGAACCAGT	AGTGACTACT	200
GGTTGTCCTT	TGTCGGTCGA	GTTAGGCCCG	GGCCTGATTT	CAGAAATGTT	250
TGACGGTATT	CAGCGACCGC	TTGATCGTTT	TCAAAAAGCA	ACGGACAGCG	300
ACTTTTTAAT	CCGTGGTGTG	GCTATCCCAA	GTCTTGATCG	AAAGGCTAAG	350
TGGGCATTTA	TTCCCAAGCT	AAGTGTGGGT	CAAGAAGTAG	TTGCAGGTGA	400
TATTTTAGGA	ACTGTGCAAG	AAACAGCTGT	CATTGAGCAC	CGTATCATGG	450
TTCTTTATAA	AGTTTCAGGG	ACCTTGGTGG	CTATTCATGC	AGGGGACTTC	500
ACAGTAACAG	ATACAGTTTA	TGAAATTAAG	CAGGAAGACG	GTTCCATTTA	550
CCAAGGTAGC	CTCATGCAGA	CTTGGCCAGT	TCGTCAAAGT	CGCCCTGTTG	600
CTCAAAAGCT	TATCCAGTC	GAACCTTTGG	TTACAGGTCA	ACGGGTTATT	650
GACACCTTTT	TCCCTGTTAC	AAAAGGTGGT	GCCGCTGCCG	TTCCTGGACC	700
ATTTGGGGCA	GGAAAAACAG	TTGTGCAGCA	TCAAATAGCT	AAATTTGCCA	750
ACGTTGATAT	TGTTATTTAT	GTGGTGTGTG	GGGAACGCGG	CAACGAGATG	800
ACCGACGTTT	TGAATGAGTT	TCCAGAGTTA	ATTGACCCAA	ATACAGGCCA	850
GTCCATTATG	GAGCGCACGG	TGTTAATTGC	AAACACCTCT	AATATGCCAG	900
TAGCAGCGCG	TGAAGCGTCG	ATTTACACAG	GTATTACCAT	TGCCGAATAT	950
TTCCGTGATA	TGGGCTATTC	TGTGGCTATC	ATGGCAGACT	CGACATCACG	1000
TTGGGCAGAA	GCTCTGCGCG	AGATGTCAGG	ACGCCTACAA	GAAATGCCTG	1050
GTGATGAAGG	CTACCCGGCT	TACTTAGGGA	GTCGTATTGC	CGAATATTAT	1100
GAACGGGCTG	GTCGTGTTTCG	GACCTTGGA	AGTCAAGAAC	GTGAGGGAAC	1150
CATTACAGCC	ATCGGCGCGG	TTTCTCCTCC	TGGAGGGGAT	ATTTTCAGAGC	1200
CTGTCACTCA	AAACACCCTT	CGGATTGTCA	AAGTTTTCTG	GGGGCTCGAC	1250
GCGCCTCTTG	CGCAACGGCG	TCACTTCCCA	GCGATTAAct	GGCTGACGTC	1300
TTATTTCATTG	TATCAAGATG	ATGTAGGAAG	CTATATTGAC	CGTAAACAGC	1350
AATCTAATTG	GTCCAACAAG	GTAACTCGTG	CCATGGCTAT	TTTGCAGCGT	1400
GAAGCCAGTC	TAGAAGAAAT	TGTACGCTTG	GTGGGGCTTG	ATTCACTGTC	1450
TGAACAAGAT	CGTTTGACCA	TGGCTGTTGC	CCGGCAAATT	CGGGAGGATT	1500
ATCTCCAGCA	AAATGCCTTT	GATTCGGTGG	ATACCTTTAC	TTCCTTTCCG	1550
AAACAAGAGG	CCATGCTAAC	CAATATTTTG	ACCTTTAATG	AGGAAGCCAG	1600
CAAAGCCCTT	TCTTTGGGAG	CTTATTTTAA	TGAGATTATG	GAAGGCACTG	1650
CTCAGGTACG	CGATCGCATC	GCACGCAGCA	AATTTATCCC	AGAAGAAAAC	1700
TTAGAGCAGA	TTAAAGGGCT	TACTCAGAAG	GTTACCAAAG	AGATTACCA	1750
CGTTTTAGCA	AAGGGAGGAA	TTTAGATGAG	CGTTCT		1786

## 60 2) INFORMATION FOR SEQ ID NO: 834

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 499 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Babesia bovis*  
 (B) STRAIN: Suarez-3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 834

15	TATCTCACGT	AAGTTTTTGC	GCGGCGGTTA	TATATCACTC	CAGGCCCTGG	50
	CTAAGTACGC	TAATACTGAC	GTTACTGTCT	ATGTGGGATG	TGGAGAGCGT	100
	GGAAACGAGA	TTGCCGAGGT	GCTTAAGGAG	TTCCCTGAGC	TGAAGACCAA	150
	GGTTGATGGC	AAGGAAGTGA	GCATTATGAA	ACGCACTTGC	TTGGTGGCCA	200
	ATACTTCAAA	CATGCCAGTG	GCCGCCAGGG	AGGCTAGTAT	CTACACTGGC	250
20	ATTACCCTAT	GTGAATACTT	CAGGGATATG	GGATACAACG	CCTGTGTGAT	300
	GGCGGATTCC	ACCAGTCGTT	GGGCTGAGGC	TTTGCCTGAG	ATATCAGGTC	350
	GTTTAGCTGA	GATGCCTGCT	GATTCAGGTT	ATCCCGCCTA	CCTTGCTTCT	400
	AGGCTTTCGG	CGTTCTATGA	GCGTGCTGGT	ACAGCTGAGT	GTATTGGAAC	450
25	ACCACTTCGT	GAAGGTTTCA	TTACCATTGT	TGGTGCTGTA	TCTCCACCA	499

2) INFORMATION FOR SEQ ID NO: 835

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 464 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Cryptosporidium parvum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 835

45	TTTCTCAAGC	TTTGAGTAAA	TATAGTAACT	CTGATGTTAT	TATTTACATT	50
	GGTTGTGGAG	AAAGAGGAAA	TGAAATGGCA	GAAGTTCTTA	CAGAATTCCT	100
	TGAGCTTTAT	ACTATGGTTG	ATGGAAAGAA	GGAGTCAATT	ATGCAAAGAA	150
	CTTGTTTAGT	AGCTAATACA	TCAAATATGC	CTGTCGCTGC	TAGAGAAGCT	200
	TCCATCTACA	CTGGTATTAC	ACTTTCTGAA	TACTTTAGAG	ATATGGGATG	250
	TAATGTTTTCT	ATGATGGCAG	ATTCAACTTC	TCGTTGGGCT	GAAGCTCTTA	300
	GAGAAATTTT	TGGTAGATTA	GCTGAAATGC	CTGCAGATTC	GGGTTACCCA	350
50	GCATATTTAG	GCGCCAGACT	TGCTTCATTC	TATGAAAGAT	CAGGAAGAGT	400
	TAAATGTATG	GGTTCCCCAG	ATAGAGAAGG	TACAGTAACA	ATTGTTGGTG	450
	CAGTTTCTCC	ACCT				464

55

2) INFORMATION FOR SEQ ID NO: 836

- 3) (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 446 bases

60

670

- (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania infantum*  
 (B) STRAIN: MOU

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 836

5 GCCCTCTCCA AGTACTCCAA CTCCGATTGC GTCATCTATG TCGGCTGCGG 50  
 CGAGCGCGGT AATGAGATGG CCGAGGTGCT CATGGAGTTC CCGACCCTGA 100  
 CGACCGTGAT CGATGGCCGC GAGGAGTCGA TCATGAAGCG CACCTGCCTC 150  
 GTGGCGAACA CCTCGAACAT GCCAGTCGCA GCGCGTGAGG CCTCTATTTA 200  
 CACCGGCATC ACCCTGGCCG AGTACTACCG TCGATATGGGC AAGCATATCG 250  
 CCATGATGGC TGACTCGACG TCTCGCTGGG CCGAGGCGCT TCGTGAGATT 300  
 TCGGGTCGTC TGGCGGAGAT GCCGGCGGAT GGTGGCTACC CCGCCTACCT 350  
 20 CAGCGCTCGT CTCGCCTCCT TCTACGAGCG CGCCGGCCTC GTCACCTGCA 400  
 TCGGCGGGCC GAAGCGCCAG GGCTCCGTCA CGATCGTCGG TGCCGT 446

25 2) INFORMATION FOR SEQ ID NO: 837

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 456 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania major*  
 (B) STRAIN: ATCC 50122

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 837

40 TTAGTCAGGC CCTCTCCAAG TACTCCAAC CCGACTGCGT CATCTATGTC 50  
 GGCTGCGGCG AGCGCGGTAA TGAGATGGCT GAGGTGCTCA TGGATTTCCTC 100  
 AACTCTGACG ACCGTGATCG ATGGTCGCGA GGAGTCCATC ATGAAGCGCA 150  
 CCTGCCTCGT GGCAAACACT TCGAACATGC CAGTCGCAGC CCGCGAGGCC 200  
 TCTATTTACA CCGGCATCAC CCTGGCCGAG TACTACCGTG ATATGGGCAA 250  
 45 GCATATTGCC ATGATGGCCG ACTCGACATC TCGCTGGGCC GAGGCGCTTC 300  
 GTGAGATTTC CGGTCGTCTG GCGGAGATGC CAGCCGATGG TGGCTACCCT 350  
 GCCTACCTCA GCGCTCGTCT CGCCTCCTTC TACGAGCGCG CCGGCCTCGT 400  
 CACCTGCATC GGCGGGCCGA AGCGCCAGGG CTCCGTACAG ATCGTCGGTG 450  
 50 CTGTGT 456

2) INFORMATION FOR SEQ ID NO: 838

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 450 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

671

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania tarentolae*  
(B) STRAIN: MOU-2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 838

10	AGGCCCTCTC	CAAGTACTCC	AACTCCGACT	GCGTCATCTA	CGTCGGCTGC	50
	GGCGAGCGCG	GTAATGAGAT	GGCCGAGGTG	CTCATGGAGT	TCCCGACCCT	100
	GACGACTGTG	ATTGATGGCC	GTGAGGAGTC	GATCATGAAG	CGGACCTGCC	150
	TCGTGGCCAA	CACCTCCAAC	ATGCCAGTCG	CAGCCCGTGA	AGCCTCTATT	200
	TACACTGGTA	TCACCCTGGC	CGAATACTAC	CGTGATATGG	GCAAGCATAT	250
	CGCCATGATG	GCTGACTCGA	CGTCTCGCTG	GGCGGAGGCG	CTTCGTGAGA	300
15	TTTCGGGTCG	CCTGGCGGAG	ATGCCGGCTG	ATGGTGGGTA	CCCCGCCTAC	350
	CTTAGTGCTC	GTCTTGCTC	CTTCTACGAG	CGTGCCGGCC	TCGTCACCTG	400
	CATCGGTGGG	CCGAAGCGCC	AGGGCTCCGT	CACGATCGTC	GGTGCCGTGT	450

20

2) INFORMATION FOR SEQ ID NO: 839

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 437 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma brucei*  
(B) STRAIN: EATRO 795

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 839

40	GCACTCTCGA	AGTACTCGAA	CAGTGACGCT	GTTATTTACG	TGGGTTGTGG	50
	TGAGCGTGCG	AATGAGATGG	CGGAGGTGCT	CATGGACTTT	CCCACCCTCA	100
	CCACCATTTAT	TGATGGACGT	GAGGAGTCTA	TCATGAAGCG	CACATGCCTG	150
	GTGGCAAATA	CTTCCAATAT	GCCTGTTGCT	GCTCGTGAGG	CATCTATTTA	200
	CACTGGTATC	ACCTTAGCTG	AGTATTATCG	GGATATGGGA	AAACACATCG	250
	CCATGATGGC	CGATTCAACC	TCCCGCTGGG	CTGAGGCTTT	GCGTGAGATT	300
	TCCGGTCGTC	TTGCTGAAAT	GCCTGCAGAT	GGAGGTTATC	CCGCGTACCT	350
45	CAGCGCCCGT	TTGGCCTCCT	TCTACGAGCG	TGCTGGCCGT	GTGACATGCA	400
	TCCGTGGGCC	GAAGCGTGAA	GGGTCTGTAA	CAATTGT		437

50 2) INFORMATION FOR SEQ ID NO: 840

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1052 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma cruzi*  
 (B) STRAIN: MM3

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 840

```

5  TGAAGGCTGA GCGCGAGCGC GGCATCACGA TCGACATCGC GCTCTGGAAG      50
   TTCGAGTCGC CCAAGTCTGT GTTCACGATC ATCGACGCCC CCGGCCACCG      100
   CGACTTCATC AAGAACATGA TCACGGGCAC GTCTCAGGCG GACGCCGCCG      150
   TCCTTGTCAT TGCGTCATCG CAGGGTGAGT TTGAGGCGGG CATCTCGAAG      200
10  GACGGCCAGA CACGCGAGCA CGCGCTGCTC GCCTTCACGC TCGGCGTGAA      250
   GCAGATGGTT GTGTGCTGCA ACAAGATGGA CGACAAGTCG GTGAAC TTCG      300
   CCCAGGAGCG CTACGATGAG ATTGTGAAGG AGGTGTCGGC GTACCTGAAG      350
   AAGGTTGGGT ACAACGTGGA GAAGGTGCGC TTCATCCCCA TCTCCGGCTG      400
   GCAGGGCGAC AACATGATTG ACAAGTCGGA AAATATGCCG TGGTACAAGG      450
15  GCCCCACGCT GCTGGAGGCA CTCGACATGC TGGAGCCCCC GGTGCGCCCC      500
   AGCGACAAGC CGCTGCGCCT GCCGCTGCAG GACGTGTACA AGATCGGCGG      550
   TATCGGCACC GTGCCGGTCG GTCGCGTGGA GACGGGCACG ATGAAGCCCC      600
   GCGACGTGGT GACGTTTGCG CCCGCCAACG TGACGACGGA GGTGAAGTCG      650
   ATTGAGATGC ACCACGAGCA GCTGGCCGAG GCCACGCCCG GCGACAACGT      700
20  CGGCTTCAAC GTGAAGAACG TGTCCGTGAA GGACATCCGC CGTGGCAACG      750
   TGTGCGCGAA CTCGAAGAAC ACAACCCCAA AGGAGGCGGC CGACTTCACG      800
   GCGCAGGTGA TCATCCTGAA CCACCCCGGC CAGATCGGCA ACGGCTATGC      850
   GCCGGTGCTC GACTGCCACA CCTGCCACAT CGCGTGCAAG TTCGCCGAGA      900
   TCGAGTCCAA GATCGACCGC CGCTCCGGCA AGGAGCTTGA GAAGAACCCC      950
25  AAGTCGATCA AGTCCGGTGA CGCCGCCATG GTGCGCATGG TGCCGCAGAA     1000
   GCCCATGTGC GTGGAGGTGT TCAACGACTA CGCTCCTCTT GGCCGCTTTG     1050
   CC                                                                1052

```

30 2) INFORMATION FOR SEQ ID NO: 841

## (i) SEQUENCE CHARACTERISTICS:

```

35  (A) LENGTH: 1061 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
   (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

40

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trypanosoma cruzi*  
 (B) STRAIN: CGL-1

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 841

```

50  TGAAGGCTGA GCGCGAGCGC GGCATCACGA TCGACATCGC GCTCTGGAAG      50
   TTCGAGTCGC CCAAGTCTGT GTTCACGATC ATCGACGCCC CCGGCCACCG      100
   CGACTTCATC AAGAACATGA TCACGGGCAC GTCTCAGGCG GATGCCGCCG      150
   TCCTTGTCAT TGCGTCATCG CAGGGTGAGT TTGAGGCGGG CATCTCGAAG      200
   GACGGCCAGA CACGCGAGCA CGCGCTGCTC GCCTTCACGC TCGGCGTGAA      250
   GCAGATGGTT GTGTGCTGCA ACAAGATGGA CGACAAGTCG GTGAAC TTCG      300
   CCCAGGAGCG CTACGATGAG ATTGTGAAGG AGGTGTCGGC GTACCTGAAG      350
   AAGGTTGGGT ACAACGTGGA GAAGGTGCGC TTCATCCCCA TCTCCGGCTG      400
55  GCAGGGCGAC AACATGATTG ACAAGTCGGA AAATATGCCG TGGTACAAGG      450
   GCCCCACGCT GCTGGAGGCA CTCGACATGC TGGAGCCCCC GGTGCGCCCC      500
   AGCGACAAGC CGCTGCGCCT GCCGCTGCAG GATGTGTACA AGATCGGCGG      550
   TATCGGCACC GTGCCGGTCG GTCGCGTGGA GACGGGCACG ATGAAGCCCC      600
   GCGACGTGGT GACGTTTGCG CCCGCCAACG TGACGACGGA GGTGAAGTCG      650
60  ATTGAGATGC ACCACGAGCA GCTGGCCGAG GCCACGCCCG GCGACAACGT      700

```



	CGGCTTCAAC	GTGAAGAACG	TGTCCGTGAA	GGACATCCGC	CGTGGCAACG	750
	TGTGCGGCAA	CTCGAAGAAC	GATCCCCCAA	AGGAGGCGGC	CGACTTCACG	800
	GCGCAGGTGA	TCATCCTGAA	CCACCCCGGC	CAGATCGGCA	ACGGCTATGC	850
	GCCGGTGCTT	GACTGCCACA	CCTGCCACAT	CGCGTGCAAG	TTCGCCGAGA	900
5	TCGAGTCCAA	GATCGACCGC	CGCTCCGGCA	AGGAGCTTGA	GAAGAACCCC	950
	AAGTCGATCA	AGTCCGGTGA	CGCCGCCATG	GTGCGCATGG	TGCCGCAGAA	1000
	GCCCATGTGC	GTGGAGGTGT	TCAACGACTA	CGCTCCTCTT	GGCCGCTTTG	1050
	CCGTGCGTGA	C				1061

10

## 2) INFORMATION FOR SEQ ID NO: 842

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 1062 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## 20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma cruzi*  
(B) STRAIN: PCU-1

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 842

	GCTGAAGGCT	GAGCGCGAGC	GCGGCATCAC	GATCGACATC	GCGCTCTGGA	50
	AGTTTCGAGTC	GCCCAAGTCT	GTGTTACAGA	TCATCGACGC	CCCCGGCCAC	100
30	CGCGACTTCA	TCAAGAACAT	GATCACGGGC	ACGTCTCAGG	CGGACGCCGC	150
	CGTCCTTGTC	ATTGCGTCAT	CGCAGGGTGA	GTTTGAGGCG	GGCATCTCGA	200
	AGGACGGCCA	GACACGCGAG	CACGCGCTGC	TCGCCTTCAC	GCTCGGCGTG	250
	AAGCAGATGG	TTGTGTGCTG	CAACAAGATG	GACGACAAGT	CGGTGAACCT	300
	CGCCCAGGAG	CGCTACGATG	AGATTGTGAA	GGAGGTGTCTG	GCGTACCTGA	350
35	AGAAGGTTGG	GTACAACGTG	GAGAAGGTGC	GCTTCATCCC	CATCTCCGGC	400
	TGGCAGGGCG	ACAACATGAT	TGACAAGTCG	GAAAATATGC	CGTGGTACAA	450
	GGGCCCCACG	CTGCTGGAGG	CACTCGACAT	GCTGGAGCCC	CCGGTGCGCC	500
	CCAGCGACAA	GCCGCTGCGC	CTGCCGCTGC	AGGACGTGTA	CAAGATCGGC	550
	GGTATCGGCA	CCGTGCCGGT	CGGTGCGCTG	GAGACGGGCA	CGATGAAGCC	600
40	CGGCGACGTG	GTGACGTTTG	CGCCCGCCAA	CGTGACGACG	GAGGTGAAGT	650
	CGATTGAGAT	GCACCACGAG	CAGCTGGCCG	AGGCCACGCC	CGGCGACAAC	700
	GTCGGCTTCA	ACGTGAAGAA	CGTGTCCGTG	AAGGACATCC	GCCGTGGCAA	750
	CGTGTGCGGC	AACTCGAAGA	ACGACCCCCC	AAAGGAGGCG	GCCGACTTCA	800
	CGGCGCAGGT	GATCATCCTG	AACCACCCCG	GCCAGATCGG	CAACGGCTAT	850
45	GCGCCGGTGC	TCGACTGCCA	CACCTGCCAC	ATCGCGTGCA	AGTTCGCCGA	900
	GATCGAGTCC	AAGATCGACC	GCCGCTCCGG	CAAGGAGCTT	GAGAAGAACC	950
	CCAAGTCGAT	CAAGTCCGGT	GACGCCGCCA	TGGTGCGCAT	GGTGCCCGAG	1000
	AAGCCCATGT	GCGTGGAGGT	GTTCAACGAC	TACGCTCCTC	TTGGCCGCTT	1050
	TGCCGTGCGT	GA				1062

50

## 2) INFORMATION FOR SEQ ID NO: 843

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1057 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Babesia bovis*  
(B) STRAIN: Suarez-3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 843

```

10  TACATTGACA GCTGCATTGA CCAAGGTCTG TTCGATGGGT GGTCATGGCG      50
    AGTACACTCC TTATGAAGCA ATTGACCGTG CTCCTGAGGA GCGTAAACGT      100
    GGTATTACTA TAAATTTCGAC ACATGTTGAA TATGAGACTA AGAACCGTCA      150
    TTACGGTCAC GTGGACTGTC CAGGTCACCTC TGATTATGTG AAGAACATGA      200
    TATCTGGCGC TGCTCAGATG GATGGTGCCA TATTGGTTGT TTCTTGTGTT      250
    GACGGTCCCA TGCCTCAGAC TAAGGAGCAC GTGTTGCTTG CTAAGCAGAT      300
15  TGGTGTACCT CGTTTAGTTG TGTTTTTGAA CAAGCTTGAC ATGTTAGAGG      350
    ACTCTGAGCT ATTGGAGTTG GTGGAGTTAG AGGTTCTGTA GTTATTGAGT      400
    GAGTTTGGTT ACGACGGTGA CAACACGCCT ATCGTTCGTG GCAGTGCTAT      450
    AAAGGCATTG AACAGTAGTT CCGAGGCTGA CATTAAAGCCA ATTCAGGATT      500
    TATTGGATGC GTGTGATGCC TTTTCTACTGA CTCCAGAACG TAAGGATGAC      550
20  ATGCCGCTCT TGGTTGCTAT TGACGATGTT CTTGCCATTC CTGGCAAGGG      600
    TACTGTTGTA ACCGGTAGGA TAGAGCAGGG CAAGATTCGT TGTGGTGACC      650
    CTATTGAGGT TTGCGCCGGT CCGAAGTCCG GCAAGAAGAC TGTGTGTGTT      700
    GGTCTTRARA TGTTCCGCAA GAGTCTCAGT RAGGGTATTG CTGGTGACCA      750
    GATTGGTGTT TTGCTCAAGG GTGTGAAGCG CGACRAGGTA GAGCGCGGGT      800
25  TTGTATTGAT TCAACCCGGA AGTTACAAAT GTCACGGTGA ATTTGATGCT      850
    GACTTGTAACG TGTTGACTAC GRAGGAAGGT GGGCGCAAGC ATCCGTTTGT      900
    GTCTAACTAC CGTCCTCAGG CGTTTATACG TACTGGAGAC GTTTGCTGCT      950
    CAGTTCATTT GGATRAGGGT GTTGAGATGG CAGCTCCTGG TGACAACGTG      1000
    CGTTGCAAGA TCAAGTTACT TTATCCCATG CCTGTCCATG AAGGTTTACG      1050
30  ATTTGCG      1057

```

2) INFORMATION FOR SEQ ID NO: 844

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 943 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Leishmania aethiopica*  
(B) STRAIN: ATCC 50119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 844

```

50  CGGTGGCATC ATTGTGGTGG CGGCCACCGA CGGCGTCATG CCGCAGACAC      50
    GCGAGCACCT CCTGATCTGC TCGCAGATTG GGCTTCCGGC GCTCGTAGGG      100
    TTCATCAACA AAGTGGATAT GACGGACGAG GACACGTGCG ACCTGGTGGA      150
    CATGGAGGTG CGCGAGCAGC TGGAGAAATA CAAGTTTCCG GCGGAGGAGA      200
    CACCAATCGT GCGCGGCTCG GCCCTCAAGG CCGTCGAGGG CGACGCGAAG      250
55  TACGAGGAGA ACATCCTCGA ACTGGTGCGG AAGTGCGACG AGTGATCCC      300
    CGACCCGCCG CGCAACACAG ACAAGCCTTT CCTTATGGCC ATCGAGCACG      350
    TTTACGAGAT CGGCAAGGAC AAGAAGAGCG TCATCGTGAC CGGCCGCGTC      400
    GATCAGGGCG TGCTGAAGCT CAACACAGAC GCCGAGCTGG CCGGCTTCAG      450
    CGCCAAGAAG TCGACGGTGA GGGTGACGGG CATCGAGATG TACCACAAGA      500
60  CGCTGAGCGA GTGCATGCCC GGTGACTCCG TCGGCGTCAG CATTGTCGGC      550

```

```

ACCGGCGACA CAACCAGTCT GTCCAAGGAC AACGTGGAAC GCGGCATGGT 600
AATGGCGGCG ACGGGTAGCA CGAACCTGTA CAACAAGGTG AAGGCGCAGG 650
TGTACGTGCT GACGAAGGAT GAGGGCGGCC GCCACACCGG CTTCAGCCCC 700
CACTACCGCC CGCAGCTCTT CTTCCATTGC GCTGACGTGA CAGCGGACAT 750
5 GAGCTTCCCG GAGGCGGAGA AGCACC GGGA GGAGCTGAAC AAGAAATTCG 800
GCCGCGGCCC CGAGGAGGAC AAGAAGAAAG AGGCGGAGAT GAAGGAGTTC 850
GAGAGCAAGC TCGTCTGCAT GCCGGGCGAT AACCGCGAGC TGATCCTGAC 900
GCTCGCGTAC CCGATGCCCA TTGAAAAGGG CCTGAAGTTC ACC 943

```

10

## 2) INFORMATION FOR SEQ ID NO: 845

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 939 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: Genomic DNA

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Leishmania amazonensis
        (B) STRAIN: ATCC 50131

```

```

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 845

```

```

CGGTGGCATC ATTGTGGTGG CGGCCACCGA CGGCGTCATG CCGCAGACAC 50
GCGAGCATCT CTTGATCTGC TCGCAGATTG GGCTTCCGGC GCTCGTAGGG 100
30 TTCATCAACA AAGTGGACAT GACGGACGAG GACACGTGCG ACCTGGTGGA 150
CATGGAGGTG CGTGAGCAGC TGGAGAAATA CAAGTTTCCG GCGGAAGAGA 200
CGCCCATCGT GCGCGGCTCG GCCCTCAAAG CCGTCGAGGG CGACGCGAAG 250
TACGAGGAGA ACATCCTCGA ACTGGTGCGG AAGTGCGACG AATGGATCCC 300
CGACCCGCCG CGCAACACAG ACAAGCCTTT CCTTATGGCC ATTGAGCAGC 350
35 TTTACGAGAT CGGCAAGGAC AAGAAGAGCG TCATCGTGAC CGGCCGCGTC 400
GATCAGGGCG TGCTGAAGCT CAACACAGAG GCCGAGCTGG CCGGCTTCAG 450
CGCCAAGAAG TCGACGGTGA GGGTGACGGG CATCGAGATG TACCACAAGA 500
CGCTGAGTGA GTGCATGCCC GGTGACTCCG TCGGCGTCAG CATTGTCGGC 550
ACCGGCGACA CGATCAGTCT CTCCAAGGAC AACGTGGAAC GCGGCATGGT 600
40 AATGGCGGCA ACGGGTAGCA CGAACCTGTA CAACAAGGTG AAGGCGCAGG 650
TGTACGTGCT GACGAAGGAT GAGGGCGGCC GCCACACTGG CTTCAGCCCC 700
CACTACCGCC CGCAGCTCTT CTTCCATTGT GCTGACGTGA CAGCGGACAT 750
GAGCTTCCCG GAGGCGGAGA AGCACC GGGA GGAGCTCAAC AAGAAATTCG 800
GCCGCGGCCC CGAGGAGGAC AAGCAGAAGG AGGCGGAGAT GAAAGAGTTC 850
45 GAGAGCAAGC TCGTCTGCAT GCCGGGCGAC AACCGCGAGC TGATCCTGAC 900
GCTGGCGTAC CCGATGCCCA TTGACAAGGG TCTGAAGTT 939

```

## 50 2) INFORMATION FOR SEQ ID NO: 846

```

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 945 bases
        (B) TYPE: Nucleic acid
55 (C) STRANDEDNESS: Double
        (D) TOPOLOGY: Linear

    (ii) MOLECULE TYPE: Genomic DNA

60 (vi) ORIGINAL SOURCE:

```

(A) ORGANISM: *Leishmania donovani*  
 (B) STRAIN: ATCC 50212

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 846

```

5  CATTGTGGTG GCGGCCACCG ACGGCGTCAT GCCGCAGACA CGCGAGCACC 50
   TCCTGATCTG CTCGCAGATC GGGCTTCCGG CGCTCGTAGG GTTCATCAAC 100
   AAGGTGGACA TGACGGACGA GGACACGTGC GACCTGGTGG ACATGGAGCT 150
   GCGCGAGCAG CTGGAGAAAT ACAAGTTTCC GGCGGAGGAG ACGCCAATCG 200
10  TGC GCGGCTC AGCCCTCAAA GCCGTCGAGG GCGATGCGAA GTACGAGGAG 250
   AACATTCTCG AACTGGTGCG GAAGTGTGAC GAGTGGATCC CTGACCCGCC 300
   GCGCAACACA GACAAGCCTT TCCTTATGGC CATCGAGCAC GTTTACGAGA 350
   TCGGCAAGGA CAAGAAGAGC GTTGTCTGTA CCGGCCGCGT CGATCAGGGC 400
   ATTCTGAAGC TCAACACAGA CGCCGAGCTG GCCGGCTTCA GCTCCAAGAA 450
15  GTCGACGGTG AGGGTGACGG GCATCGAGAT GTACCACAAG ACGCTGAGCG 500
   AGTGCATGCC TGGTGACTCC GTCGGCGTCA GCATTGTCGG GACCGGCGAC 550
   ACGACCAGTC TATCCAAGGG CAACGTGGAA CGCGGCATGG TGATGGCGGC 600
   GACGGGTAGC ACGAACCTGT ACAACAAGGT GAAGGCGCAG GTGTACGTGC 650
   TGACGAAGGA TGAGGGCGGC CGCCACACTG GCTTTAGTCC TCACTACCGC 700
20  CCGCAGCTCT TCTTCCATTG TGCTGACGTG ACGGCGGACA TGAGCTTCCC 750
   GGAGGCGGAG AAGCACCGCG AAGAGCTCAA CAAGAAATTC GGCCGCGGCC 800
   CCGAGGAGGA CAAGAAGAAA GAGGCAGCGA TGAAGGAGTT CGAGAGCAAG 850
   CTCGTCTGCA TGCCGGGCGA TAACCGCGAG CTGATCCTGA CGCTGGCGTA 900
   CCCGATGCCC ATTGAAAAGG GTCTGAAGTT CACCATCCGT GAGGG 945
25

```

## 2) INFORMATION FOR SEQ ID NO: 847

```

30  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 939 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear
35
      (ii) MOLECULE TYPE: Genomic DNA

      (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Leishmania infantum
          (B) STRAIN: MOU
40

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 847

```

45  GGCATCATTG TGGTGGCGGC CACCGACGGC GTCATGCCGC AGACACGCGA 50
   GCACCTCCTG ATCTGCTCGC AGATCGGGCT TCCGGCGCTC GTAGGGTTCA 100
   TCAACAAGGT GGACATGACG GACGAGGACA CGTGCGACCT GGTGGACATG 150
   GAGCTGCGCG AGCAGCTGGA GAAATACAAG TTTCCGGCGG AGGAGACGCC 200
   AATCGTGCGC GGCTCAGCCC TCAAAGCCGT CAGGGCGGAT CGGAAGTACG 250
   AGGAGAACAT TCTCGAACTG GTGCGGAAGT GTGACGAGTG GATCCCTGAC 300
50  CCGCCGCGCA ACACAGACAA GCCTTTCCTT ATGGCCATCG AGCACGTTTA 350
   CGAGATCGGC AAGGACAAGA AGAGCGTTGT CGTGACCGGC CGCGTCGATC 400
   AGGGCGTTCT GAAGCTCAAC ACAGACGCCG AGCTGGCCGG CTTCAGCTCC 450
   AAGAAGTCGA CGGTGAGGGT GACGGGCATC GAGATGTACC ACAAGACGCT 500
   GAGCGAGTGC ATGCCTGGTG ACTCCGTCGG CGTCAGCATT GTCGGCACCG 550
55  GCGCAGCAGC CAGTCTATCC AAGGCCAACG TGGAAACGCG CATGGTGATG 600
   GCGGCGCAGG GTAGCACGAA CCTGTACAAC AAGGTGAAGG CGCAGGTGTA 650
   CGTGCTGACG AAGGATGAGG GCGGCCGCCA CACTGGCTTT AGTCCTCACT 700
   ACCGCCCGCA GCTCTTCTTC CATTGTGCTG ACGTGACGGC GGACATGAGC 750
   TTCCCGGAGG CGGAGAAGCA CCGCGAAGAG CTCAACAAGA AATTCGGCCG 800
60  CGGCCCCGAG GAGGACAAGA AGAAAGAGGC AGCGATGAAG GAGTTCGAGA 850

```

GCAAGCTCGT CTGCATGCCG GGCGATAACC GCGAGCTGAT CCTGACGCTG 900  
 GCGTACCCGA TGCCCATGGA AAAGGGTCTG AAGTTCACC 939

5

## 2) INFORMATION FOR SEQ ID NO: 848

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 933 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania enriettii*  
 (B) STRAIN: ATCC 50120

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 848

25 CACCGACGGC GTCATGCCGC AGACACGGGA GCACCTGCTC ATCTGCTCGC 50  
 AGATCGGGCT GCCGGCGCTT GTAGGGTTCA TCAATAAAGT TGACATGACG 100  
 GACGAGGATA CGTGCGACCT CGTGGACATG GAGGTGCGGG AACAGCTGGA 150  
 GAAGTACAAG TTTCCGGCCG AGGAGACGCC CATCGTACGT GGCTCGGCCC 200  
 TCAAGGCCCT CGAGGGGGAT GCGCAATACG AGGGGAGTAT TCTCGAGCTG 250  
 GTGCGAAAGT GCGACGAGTG GATCCCCGAC CCGCCGCGCA ACACCGAAAA 300  
 GCCTTTCCTC ATGGCTATCG AGCACGTTTA CGAGCTCGGC AAAGACAAGA 350  
 AGAGCGTCAT CGTTACCGGC CGCGTCGATC AAGGTGTGCT GAAGCTCAAC 400  
 30 ACAGACGCCG AGCTGGCCGG CTTACGCGCC AAGAAGGCGA CAGTCAAAGT 450  
 GACGGGCATC GAGATGTATC ACAAGACACT EAATGAGTGC ATGCCCCGCG 500  
 ACTCTGTGCG TGTCAGCATC GTCGGTACCG GTGACACGAC CAGCTTATCC 550  
 AAGGATAATG TTGAGCGCGG TATGGTAATG GCGGCAACGG GTAGCACGAA 600  
 CCTGTACAAC AAGCTGAAGG CGCAGGTTTA CGTGCTGACA AAGGAGGAGG 650  
 35 GTGGCCGCCA CACCGGGTTC AGCCCCCACT ACCGCCGCA GCTCTTCTTC 700  
 CACTGCGCTG ACGTGACCGC AGACATGAGC TTCCCGGAGG CGGAGAAGTA 750  
 CCGCGAGGAG CTCAACAAGA AGTTCGGCCG TGGCCCTGAG GAGGACAAGA 800  
 AGAAAGAGGC GGAGATGAAG GAGTTCGAAA GCAAACCTGT CTGCATGCCA 850  
 GGCGATAACC GCGAGCTGAT CCTAACTCTG GCGTACCCGA TGCCCATCGA 900  
 40 CAAGGGCCTG AAGTTCACCA TCCGTGAGGG CGG 933

## 2) INFORMATION FOR SEQ ID NO: 849

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 943 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania gerbilli*  
 (B) STRAIN: ATCC 50121

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 849

60 CCGTGGCATC ATTGTGGTGG CGGCCACCGA CGGCGTCATG CCGCAGACAC 50

```

5  GCGAGCACCT CCTGATCTGC TCGCAGATTG GGCTTCCGGC GCTCGTAGGG 100
   TTCATCAACA AAGTGGACAT GACGGACGAG GACACGTGCG ACCTGGTGGA 150
   CATGGAGGTG CGCGAGCAGC TGGAGAAATA CAAGTTTCCG GCGGAGGAGA 200
   CACCAATCGT GCGCGGCTCG GCCCTCAAGG CCGTCGAGGG CGACGCGAAG 250
10 TACGAGGAGA ACATCCTCGA ACTGGTGCGG AAGTGCGACG AGTGGATCCC 300
   CGACCCGCCG CGCAACACAG ACAAGCCTTT CCTTATGGCC ATCGAGCACG 350
   TTTACGAGAT CGGCAAGGAC AAGAAGAGCG TCATCGTGAC CGGCCGCGTC 400
   GATCAGGGCG TGCTGAAGCT CAACACGGAC GCCGAGCTGG CCGGCTTCAG 450
   CGCCAAGAAG TCGACGGTGA GGGTGACGGG CATTGAGATG TACCACAAGA 500
15 CGCTGAGCGA GTGCATGCCC GGTGACTCCG TCGGCGTCAG CATTGTCTGGC 550
   ACCGGCGACA CGACCACTCT GTCCAAGGAC AACGTGGAAC GCGGCATGGT 600
   AATGGCGGCG ACGGGTAGCA CGAACCTGTA CAACAAGGTG AAGGCGCAGG 650
   TGTACGTGCT GACGAAGGAT GAGGGCGGCC GCCACACTGG CTTAGCCCC 700
   CACTACCGCC CGCAGCTCTT CTTCCATTGC GCTGACGTGA CAGCGGACAT 750
15 GAGCTTCCCG GAGGCGGAGA AGCACC GCGA GGAGCTCAAC AAGAAATTCG 800
   GCCGCGGCCG CGAGGAGGAC AAGAAGAAAG AGGCGGAGAT GAAGGAGTTC 850
   GAGAGCAAGC TCGTCTGCAT GCCGGGCGAT AACC GCGAGC TGATCCTGAC 900
   GCTGGCGTAC CCGATGCCCC TTGAAAAGGG TCTGAAGTTC ACC 943

```

20

## 2) INFORMATION FOR SEQ ID NO: 850

- (i) SEQUENCE CHARACTERISTICS:
- 25 (A) LENGTH: 918 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 30 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Leishmania major*  
 (B) STRAIN: ATCC 50122
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 850

```

40 GGCATCATTG TGGTGGCGGC CACCGACGGC GTCATGCCGC AGACACGCGA 50
   ACACCTCCTG ATCTGCTCGC AAATTGGCCT TCCGGCGCTC GTAGGGTTCA 100
   TCAACAAAGT GGACATGACG GACGAGGACA CGTGTGACCT GGTGGACATG 150
   GAGGTGCGCG AGCAGCTGGA GAAATACAAG TTTCCGGCGG AGGAGACACC 200
   AATCGTGCGC GGCTCGGCC TCAAGGCCGT CGAGGGCGAC GCGAAGTACG 250
   AGGAGAACAT CCTCGAACTG GTGCGGAAGT GCGACGAGTG GATCCCCGAC 300
   CCGCCGCGCA ACACAGACAA GCCTTTCCTT ATGGCCATCG AGCACGTTTA 350
45 CGAGATCGGC AAGGACAAGA AGAGCGTCAT CGTGACCGGC CGCGTCGATC 400
   AGGGCGTGCT GAAGCTCAAC ACAGACGCCG AGCTGGCCGG CTTAGCGGCC 450
   AAGAAGTCGA CGGTGAGGGT GACGGGCATT GAAATGTACC ACAAGACGCT 500
   GAGCGAGTGC ATGCCCAGGT ACTCCGTCGG GCTCAGCATT GTCGGCACCG 550
   GCGACACGAC CAGTCTGTCC AAGGACAACG TGGAGCGCGG CATGGTAATG 600
50 GCGGCGACGG GTAGCACGAA CCTGTACAAC AAGGTGAAGG CGCAGGTGTA 650
   CGTGCTGACG AAGGATGAGG GCGGCCGCCA CACTGGCTTC AGCCCCCACT 700
   ACCGCCCGCA GCTCTTCTTC CATTGCGCTG ACGTGACAGC GGACATGAGC 750
   TTCCCGGAGG CGGAGAAGCA CCGCGAGGAG CTCAACAAGA AATTCGGCCG 800
   CGGCCCCGAG GAGGACAAGA AGAAAGAGGC GGAGATGAAG GAGTTCGAGA 850
55 GCAAGCTCGT CTGCATGCCG GGCATAACC GCGAGCTGAT CCTGACGCTG 900
   GCGTACCCGA TGCCCATT 918

```

## 60 2) INFORMATION FOR SEQ ID NO: 851

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 939 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania mexicana*  
 (B) STRAIN: ATCC 50156

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 851

```

15  CCGTGGCATC ATTGTGGTGG CGGCCACCGA CGGCGTCATG CCGCAGACAC      50
    GCGAGCATCT CCTGATCTGC TCGCAGATTG GGCTTCCGGC GCTCGTAGGG      100
    TTCATCAACA AAGTGGACAT GACGGACGAG GACACGTGCG ACCTGGTGGA      150
    CATGGAGGTG CGTGAGCAGC TGGAGAAATA CAAGTTTCCG GCGGAAGAGA      200
20  CGCCCATCGT GCGCGGCTCG GCCCTCAAGG CCGTCGAGGG CGACGCGAAG      250
    TACGAGGAGA ACATCCTCGA ACTGGTGCGG AAGTGCACG AATGGATCCC      300
    CGACCCGCCG CGCAACACAG ACAAGCCTTT CCTTATGGCC ATTGAGCACG      350
    TTTACGAGAT CGGCAAGGAC AAGAAGAGCG TCATCGTGAC CGGCCGCGTC      400
    GATCAGGGCG TGCTGAAGCT CAACACAGAC GCCGAGCTGG CCGGCTTCAG      450
25  CGTCAAGAAAG TCGACGGTGA GGGTGACGGG CATCGAGATG TACCACAAGA      500
    CGCTGAGTGA GTGCATGCCC GGTGACTCCG TCGGCGTCAG CATTGTGCGC      550
    ACCGGCGACA CGATCAGTCT CTCCAAGGAC AACGTTGAAC GCGGCATGGT      600
    AATGGCGGCA ACGGGTAGCA CGAACCCTGTA CAACAAGGTG AAGGCGCAGG      650
    TGTACGTGCT GACGAAGGAT GAGGGCGGCC GCCACACTGG CTTCAGCCCC      700
30  CACTACCGCC CGCAGCTCTT CTTCCATTGT GCTGACGTGA CGGCGGACAT      750
    GAGCTTCCCG GAGGCGGAGA AGCACCGCGA GGAGCTCAAC AAGAAATTCG      800
    GCCGCGGCCG CGAGGAGGAC AAGCAGAAGG AGGCGGAGAT GAAAGAGTTC      850
    GAGAGCAAGC TCGTCTGCAT GCCGGGCGAC AACC GCGAGC TGATCCTGAC      900
    GCTGGCGTAC CCGATGCCCA TTGAGAAGGG TCTGAAGTT      939
35

```

## 2) INFORMATION FOR SEQ ID NO: 852

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 912 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania tarentolae*  
 (B) STRAIN: MOU-2

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 852

```

55  TCATTGTGGT GGCCGCCACC GACGGCGTCA TGCCGCAAAC ACGGGAGCAC      50
    CTTTGTATCT GCTCGCAGAT CGGGCTGCCG GCGCTCGTAG GGTTCATCAA      100
    CAAAGTGGAC ATGACAGACG AAGACACGTG CGACCTGGTA GACCTGGAGG      150
    TGCGTGAGCA GCTGGAGAAG TACAAGTTTC CGGCAGAGGA AACACCAATC      200
    GTGCGTGGCT CGGCCCTCAA GGCCGTTGAG GGCGATGCAA AGTACGAGGA      250
    GAACATCCTC GAACTGGTGC GGAAGTGCGA CGAGTGGATC CCAGACCCGC      300
60  CACGCAATAC GGACAAGCCT TTCCTTATGG CCATTGAACA CGTGTACGAG      350

```

	ATCGGCAAGG	ATAGGAAAAG	CGTCATCGTA	ACCGGCCGCG	TCGATCAAGG	400
	TGTGCTGAAG	CTGAACACAG	ACGCCGAGCT	GGCCGGCTTC	AGCGCCAAGA	450
	AGTCGACGGT	GAAAGTGACG	GGCATTGAGA	TGTACCACAA	GACACTGACA	500
	GAGTGCATGC	CCGGCGACTC	TGTCGGCGTC	AGCATTGTGG	GCACTGGYGA	550
5	CACGACCAGC	CTCTCTAAGG	ACAATGTTGA	GCGTGGCATG	GTA CTGGCCG	600
	CTACGGGTAG	CACGAACCTG	TACAACAAAG	TAAAGGCGCA	GGTGTATGTA	650
	CTCACGAAGG	ATGAGGGCGG	CCGCCACACC	GGCTTCAGCC	CCCACTACCG	700
	TCCGCAGCTC	TTCTTCCACT	GCGCTGACGT	AACGGCGGAC	ATGAGCTTCC	750
	CGGAGGCGGA	GAAGCACCGC	GAGGAACTCA	ATAAGAAATT	CGGCCGCGGC	800
10	CCCGAGGAGG	ACAAGAAAAA	GGAGGCGGAG	ATGAAGGAGT	TCGAGAGCAA	850
	GCTGGTCTGC	ATGCCAGGCG	ATAACCGCGA	GCTGATCCTG	ACATTGGCGT	900
	ACCCGATGCC	TA				912

15

## 2) INFORMATION FOR SEQ ID NO: 853

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 936 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Trypanosoma cruzi*
  - (B) STRAIN: MM3

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 853

	ATTCTTGTGG	TGGCAGCTAA	CGACGGATGC	ATGCCGCAGA	CGCGTGAGCA	50
	CCTGCTTATT	TGTTTCGAGA	TTGGCCTTCC	TGCTCTTGTA	TGCTTTATCA	100
	ATAAGTGTGA	CATGATGCAA	GGGCAGGAGG	AAATGATTGA	ACTTGTTGAA	150
35	ATGGAGGTAC	GTGAACTTTT	GGAGAAGTAC	AAGTTCCCTG	CGGAGGAGAC	200
	GCCATTTGTG	CGGGGGTCTG	CGGTGAAGGC	ATTGGAGGGT	GATGCTGAAA	250
	ATGAAGGAAA	GATTTTGGAG	CTTGTA AAAA	AATGTGATGA	ATGGATTCCC	300
	GACCCACCGC	GTGCCATTGA	AAAACCGTTC	CTTATGGCCA	TTGAGCACGT	350
	TTTTGAGGTT	GGAAAGGATA	AGAAGGCCGT	TGTTGTGAGC	GGGCGTGTGG	400
40	ACCAGGGGCA	GTTGAAGGTC	GGCGCAGATG	CAGAACTTTC	CGGGTTTAGC	450
	GCAAAGAAGC	TGACGGTGAA	GGTTGCTAGC	ATCGAAATGT	ACCATAAAAAT	500
	TCTGGAGGAT	TGCATGCCTG	GTGACTCTGT	TGGCGCGAAG	ATCGTTGGCA	550
	GCGGTGAAAC	AGTGAACCTG	TCGAAGGAAA	ATGTGGAACG	CGGCATGGTA	600
	CTCTCCGCAC	CAGGTGCAAC	GACACTGTTC	AACAAGGTCC	GCGCGCAGGT	650
45	GTACGTGTTG	ACAAAGGAAG	AAGGCGGTCT	TCACACAGCC	TTTAGTCCTC	700
	ACTATCGTCC	GCAGCTTTTC	TTCCACTGTG	CTGATGTCAC	GGCAGATATT	750
	AACTTCCCGG	AAAGCGAGAA	GCTTGCAGGG	GAGCTGAACA	AAAAGTATGG	800
	CCGTGATGCG	TCGGAACAGA	AGAAGAAGGA	GGCAGAACTG	AAAGAGTTTG	850
	AAAAGACGCT	TGTCTGCATG	CCTGGTGATA	ACCGCGAACT	CCTGCTCACC	900
50	CTTGCCTATC	CAATGCCAAT	GGAAAAGGGA	CTCAAG		936

## 2) INFORMATION FOR SEQ ID NO: 854

55

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 934 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear

60



(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma cruzi*  
(B) STRAIN: PCU-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 854

10	CGGCATTCTT	GTGGTGGCAG	CTAACGACGG	ATGCATGCCG	CAGACGCGTG	50
	AGCACCTGCT	TATTTGTTTCG	CAGATTGGCC	TTCCTGCTCT	TGTATGCTTT	100
	ATCAATAAGT	GTGACATGAT	GCAAGGGCAG	GAGGAAATGA	TTGAGCTTGT	150
	TGAAATGGAG	GTACGTGAAC	TTTTGGAGAA	GTACAAGTTC	CCTGCGGAGG	200
	AGACGCCATT	TGTGCGGGGG	TCTGCGGTGA	AGGCATTGGA	GGGTGATGCT	250
15	GAAAATGAAG	GAAAGATTTT	GGAGCTTGTA	AAAAAATGTG	ATGAATGGAT	300
	TCCCGACCCA	CCGCGTGCCA	TTGAAAAACC	GTTCCCTTATG	GCCATTGAGC	350
	ACGTTTTTGA	GGTTGGAAAG	GATAAGAAGG	CCGTTGTTGT	GAGCGGGCGT	400
	GTGGACCAAG	GGCAGTTGAA	GGTCGGCGCA	GATGCAGAAC	TTTCCGGGTT	450
	TAGTGCAAAG	AAGCTGACGG	TGAAGGTTGC	TAGCATCGAA	ATGTACCATA	500
20	AAATTCTGGA	GGATTGCATG	CCTGGTGAAT	CTATTGGCGC	GAAGATCGTT	550
	GGCAGCGGTG	AAACAGTGAA	CCTGTGCAAG	GAAAATGTGG	AACGCGGCAT	600
	GGTACTCTCC	GCACCAGGTG	CAACGACACT	GTTCAACAGG	GTCCGCGCGC	650
	AGGTGTACGT	GTTGACAAAG	GAAGAAGGCG	GTCGTCACAC	AGCCTTTAGT	700
	CCTCACTATC	GTCCGCAGCT	TTTCTTCCAC	TGTGCTGATG	TCACGGCGGA	750
25	TATTAACCTC	CCGGAAAGCG	AGAAGCTTGC	AGGGGAGCTG	AACAAAAAGT	800
	ATGGCCGTGA	CGCGGCGGAA	CAGAAGAAGA	AGGAGGCAGA	ACTGAAAGAG	850
	TTTGAAAAGA	CGCTTGCTCTG	CATGCCTGGT	GATAACCGCG	AACTCCTGCT	900
	CACCCTTGCC	TATCCAATGC	CAATGGAAAA	GGGA		934

30

2) INFORMATION FOR SEQ ID NO: 855

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 937 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma cruzi*  
(B) STRAIN: CGL-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 855

50	TGGCGGCATT	CTTGTGGTGG	CAGCTAACGA	CGGATGCATG	CCGCAGACGC	50
	GTGAGCACCT	GCTTATTTGT	TCGCAGATTG	GCCTTCCTGC	TCTTGTATGC	100
	TTTATCAATA	AGTGTGACAT	GATGCAAGGG	CAGGAGGAAA	TGATTGAACT	150
	TGTTGAAATG	GAGGTACGTG	AACTTTTGGA	GAAGTACAAG	TTCCCTGCGG	200
	AGGAGACGCC	ATTTGTGCGG	GGGTCTGCGG	TGAAGGCATT	GGAGGGTGAT	250
	GCTGAAAATG	AAGGAAAGAT	TTTGGAGCTT	GTAAAAAAT	GTGATGAATG	300
	GATTCCTCGAC	CCACCGCGTG	CCATTGAAAA	ACCGTTCCTT	ATGGCCATTG	350
55	AGCACGTTTT	TGAGGTTGGA	AAGGATAAGA	AGGCCGTTGT	TGTGAGCGGG	400
	CGTGTGGACC	AGGGGCAGTT	GAAGGTCGGC	GCAGATGCAG	AACTTTCCGG	450
	GTTTAGCGCA	AAGAAGCTGA	CGGTGAAGGT	TGCTAGCATC	GAAATGTACC	500
	ATAAAATTCT	GGAGGATTGC	ATGCCTGGTG	ACTCTGTTGG	CGCGAAGATC	550
	GTTGGCAGCG	GTGAAACAGT	GAACCTGTCTG	AAGGAAAATG	TGGAACGCGG	600
60	CATGGTACTC	TCCGCACCAG	GTGCAACGAC	ACTGTTCAAC	AAGGTCCGGC	650

	CGCAGGTGTA	CGTGTGACA	AAGGAAGAAG	GCGGTCGTCA	CACAGCCTTT	700
	AGTCCTCACT	ATCGTCCGCA	GCTTTTCTTC	CACTGTGCTG	ATGTCACGGC	750
	AGATATTAAC	TTCCCGGAAA	GCGAGAAGCT	TGCAGGGGAG	CTGAACAAAA	800
	AGTATGGCCG	TGATGCGGCG	GAACAGAAGA	AGAAGGAGGC	AGAACTGAAA	850
5	GAGTTTGAAA	AGACGCTTGT	CTGCATGCCT	GGTGATAACC	GCGAACTCCT	900
	GCTCACCTT	GCCTATCCAA	TGCCAATGGA	AAAGGGA		937

## 10 2) INFORMATION FOR SEQ ID NO: 856

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 888 bases  
 (B) TYPE: Nucleic acid  
 15 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 20 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia bigemina*  
 (B) STRAIN: Suarez-2

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 856

25	CTTATTATGG	AGCTGATCAA	CAACGTCGCG	AAGAAGCACG	GTGGTTTCTC	50
	CGTGTTCGCC	GGCGTCGGCG	AGCGCACCAG	GGAGGGCAAC	GAGCTGTACC	100
	ACGAGATGAT	GGAGACCGGC	GTCATCAAGC	GCCGCCAGCT	GGATGACGGC	150
	ACGTTTCGACT	TCTCCGGCTC	CAAGGCCGCG	CTGGTGTACG	GCCAGATGAA	200
30	CGAGCCGCCA	GGTGCCAGGG	CGCGTGTGTC	CCTCACTGGC	CTGACGGTGG	250
	CCGAGTACTT	CCGTGATGAG	GACGGCCAGG	ACGTGCTGCT	CTTCATCGAC	300
	AACATCTACC	GTTTCACCCA	GGCTGGTTCT	GAGGTGAGTG	CCCTTTTGGG	350
	GCGCATCCCG	TCCGCCGTCG	GTTACCAGCC	GACCCTCGCC	ACCGACCTTG	400
	GCGCGCTGCA	GGAGCGTATC	ACGACGACCA	ACAAGGGCTC	CATCACCTCC	450
35	GTGCAGGCCG	TCTACGTGCC	GGCCGACGAT	ATCACCGACC	CGGCGCCTGC	500
	GACCACCTTC	ACCCATCTGG	ACGCGACCAC	TGTGCTCTCC	CGTTCCATCG	550
	CCGAGCTGGG	TATCTACCCC	GCCGTCGACC	GCCTCGACTC	CACCTCGCGT	600
	ATGCTGTCCG	CGAACATCGT	CGGCGAGGAG	CAGTACAACG	TGGCGCGTGG	650
	CGTGCAGAAA	ATACTGCAGG	ACTACAAATC	GCTGCAGGAT	ATCATCGCCA	700
40	TCCTGGGTAT	GGACGAGCTG	TCTGAGCAGG	ACAAGTTTCG	CGTCGCGCGT	750
	GCGCGCAAGG	TTCAGCGTTT	CCTATCCCAG	CCCTTCCAGG	TGGCTGAGGT	800
	ATTACCGGC	AAGCCCGGAC	GTTTCGTCTG	GCTGCAGGAC	ACCATCAGCG	850
	GCGTCAAGGA	GATTTTGGAC	GGCGAGTGCG	ACGACATG		888

45

## 2) INFORMATION FOR SEQ ID NO: 857

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 884 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 50 (D) TOPOLOGY: Linear

## 55 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia bovis*  
 (B) STRAIN: Suarez-3

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 857

	TGATTATGGA	ATTGATCAAC	AATGTCGCCA	AGAAACACGG	TGGGTTCTCC	50
	GTGTTTCGCTG	GTGTTGGTGA	ACGTACGAGG	GAAGGTAACG	AACTGTACCA	100
5	TGAAATGATG	GAAACGGGTG	TCATCAAGCG	CCGTCAACTG	GAAGACGGAA	150
	CATTTGACTT	CTCGGGCTCT	AAAGCTGCTT	TGGTGTACGG	ACAAATGAAC	200
	GAACCACCAG	GTGCTAGAGC	CCGTGTTGCA	CTCACGGGAT	TGACCGTTGC	250
	CGAGTATTTT	CGTGATGAAG	AGGGGCAGGA	TGTGCTACTC	TTCATCGATA	300
	ACATCTACCG	TTTCACCCAG	GCCGGTTCCG	AAGTGAGTGC	GCTGTTAGGA	350
10	AGAATTCCAT	CCGCCGTGGG	TTATCAACCT	ACATTGGCCA	CTGATCTCGG	400
	AGCACTCCAG	GAACGCATTA	CTACAACCAA	CAAGGGTTTCG	ATTACATCAG	450
	TCCAGGCAGT	ATACGTCCCA	GCCGATGATA	TCACTGATCC	CGCTCCAGCT	500
	ACCACTTTCT	CGCACTTGGA	TGCCACTACA	GTGCTTTCTC	GTTCAATTGC	550
	GGAGTTGGGT	ATTTACCCTG	CGGTTCGACC	GCTTGACTCA	ACGTCACGTA	600
15	TGCTGTCCGC	CAACATTGTA	GGACAGGAAC	AGTACGATGC	CGCACGTGGT	650
	GTACAGAAAA	TTTTACAGGA	CTACAAATCA	CTGCAGGATA	TCATTGCCAT	700
	TCTGGGTAGT	GACGAGCTGT	CTGAGCAGGA	CAAGTTCGTT	GTAGCACGCG	750
	CCCGTAAGGT	ACAGCGTTTC	CTGTCTCAGC	CGTTCCAAGT	GGCTGAGGTG	800
	TTCACCGGCA	AGCCTGGGAG	GTTCGTTGAA	CTACAGGATA	CCATCAGCGG	850
20	TGTCAAGGAA	ATCTGGAAGG	TGAGTGTGAC	GATA		884

## 2) INFORMATION FOR SEQ ID NO: 858

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 871 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia microtti*  
 (B) STRAIN: Persing-1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 858

40	TGGAAC TGAT	TAATAATGTG	GCCAAAAAGC	ATGGCGGTTA	CTCTGTTTTT	50
	GCAGGTGTAG	GTGAAAGGAC	GAGGGAGGGT	AATGAATTGT	ACCATGAAAT	100
	GATGGAGACA	GGTGTTATAA	AGAAAAAGGC	ACTAGGTGGT	GGGAAGTTTG	150
	ATTTCA GTGG	ATCTAAAGCA	GCGCTGGTCT	ATGGACAAAT	GAACGAGCCA	200
	CCTGGGGCCC	GTGCTAGAGT	GGCACTAACT	GGATTAACAG	TCGCAGAATA	250
45	TTTCCGTGAC	GAACAAGGAC	AAGACGTGTT	GTTGTTTATT	GATAATATTT	300
	ACCGATTTAC	TCAGGCAGGG	TCTGAGGTTT	CAGCCTTGCT	AGGCCGTATA	350
	CCTTCAGCTG	TGGGATACCA	GCCTACATTG	GCAACAGATC	TTGGCTGTTT	400
	ACAAGAACGA	ATTACTACGA	CCAAATCTGG	TTCAATCACC	AGTGTAACAAG	450
	CTGTGTATGT	GCCAGCAGAT	GATATTACTG	ATCCAGCGCC	TGCCACAAC	500
50	TTTACTCACT	TGGACGCTAC	TACTGTACTT	AGCAGGCCAA	TTGCTGAACT	550
	CGGTATTTAT	CCAGCGGTAG	ACCCGTTGGA	TTCAACAAGC	CGTATGCTAA	600
	GCGCGAACAT	TGTGGGAAAT	GAACACTATA	GTGTAGCCCG	TTCCGTGCAG	650
	AAGATACTGC	AAGATTACAA	ATCGCTTCAG	GACATTATTG	CCATTTTGGG	700
	TATGGATGAA	CTGTCCGAAC	AAGACAAAAA	TATAGTAGCC	CGAGCAAGGA	750
55	AGATGCAAAG	GTTCTTATCA	CAGCCATTCC	AAGTGCGCGA	AGTTTTTACT	800
	GGTAAACCGG	GAAGATTTGT	GGAATTGGAA	GATACAATTG	CCGGGGCACG	850
	AGATATAATT	GCGGGTAATT	G			871

## 2) INFORMATION FOR SEQ ID NO: 859

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1286 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania guyanensis*  
 (B) STRAIN: ATCC 50126

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 859

```

TCTCGGAGGG CGTCCCGCCC GTGCTGACGG CGCTGGATGT GACGGAGGAC      50
CTTGCCGCG ATGAGCCGCT GACGCTGGAG ATCGTGCAGC ACCTGGACGC      100
GAACACCGGC CGCTGCATTG CGATGCAGAC GACGGACCTG CTGAAGCTGA      150
AGTCGAAGGT TGTGTCGACC GGCGGCAACA TCTCTGTGCC GGTGGGCCGT      200
GAGACGCTGG GCCGCATCTT CAACGTGCTG GGCGACGCGA TCGACCAGCG      250
CGGCCCCGTG GGTGAGAAGA TGCGCATGGC GATCCACGCC GAGGCCCCGA      300
AGCTGGCGGA TCAGGCCGCG GAGGACACGA TCCTGACGAC CGGCATCAAG      350
GTGATCGACC TGATTCTGCC CTACTGCAAG GGTGGCAAGA TCGGCCTGTT      400
TGGCGGCGCC GGTGTGGGCA AGACCGTGAT CATCATGGAG CTGATTAACA      450
ACGTCGCGAA GGGCCACGGT GGTTTCTCGG TGTTTGCCGG CGTTGGCGAG      500
CGCACGCGCG AGGGCACGGA CCTGTACCTG GAGATGATGC AGTCGAAGGT      550
GATTGACCTG AAGGGCGAGT CGAAGTGCGT GCTTGTGTAC GGGCAGATGA      600
ACGAGCCCCC GGTGCGCGCG GCGCGCGTTG CGCAGTCTGC GCTGACGATG      650
CGCGAGTACT TCCGAGACGT GGAGGGCCAG AATGTGCTGC TGTTCATCGA      700
CAACATCTTC CGCTTCACGC AGGCGAACTC CGAGGTCTCT GCGCTGCTGG      750
GCCGCATTCC GGCCGCCGTG GGCTACCAGC CGACGCTTGC GGAGGATCTT      800
GGTATGCTGC AGGAGCGCAT CACGTCGACG ACGAAGGGGT CGATCACGTC      850
CGTGCAGGCC GTGTACGTGC CTGCGGATGA TATCACGGAT CCGGCGCCCCG      900
CGACGACGTT CTCGCACCTG GACGCGACGA CTGTGCTGGA CCGCGCGGTG      950
GCGGAGTCGG GGATCTACCC TGCCGTGAAC CCGCTGGAGT GCGCGTCGCG     1000
TATCATGGAC CCCGATGTGA TCGACGTGGA CCACTACAAC GTTGCGCAGG     1050
ATATCGTGCA GATGCTGACC AAGTACAAGG AGCTGCAGGA TATCATTGCG     1100
GTGCTTGTA TCGACGAGCT GAGCGAGGAG GACAAGGTCG TGGTGGACCG     1150
CGCGCGCAAG GTGACCCGGT TCCTGTCGCA GCCGTTCAG GTTGCGGAGG     1200
TGTTACGGG CATGACGGGC CACTACGTGC AGCTGGCCGA CACGGTGGAG     1250
TCGTTCTCTG GGCTGCTGAT GGGGTCGTAC GACCAG                      1286

```

## 2) INFORMATION FOR SEQ ID NO: 860

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1222 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania mexicana*  
 (B) STRAIN: ATCC 50156

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 860

```

CTCGGAGGGC GTCCCGCCCC TGCTGACGGC GCTGGATGTG ACGGAGGACC 50
TTGGCCGCGA TGAGCCGCTG ACGCTGGAGA TCGTGACGCA CCTGGACGCG 100
AACACCGGCC GCTGCATTGC GATGCAGACG ACGGACCTGC TGAAGCTGAA 150
5  GTCGAAGGTT GTGTGACCCG GCGGCAACAT CTCTGTGCCG GTGGGCGGTG 200
AGACGCTGGG CCGCATCTTC AACGTGCTGG GCGACGCGAT CGACCAGCGC 250
GGCCCCGTGG GTGAGAAGAT GCGCATGGCG ATCCACGCCG AGGCCCCGAA 300
GCTGGCGGAT CAGGCCGCGG AGGACACGAT CCTGACGACC GGCATCAAGG 350
TGATCGACCT GATTCTGCCC TACTGCAAGG GTGGCAAGAT CGGCCTGTTT 400
10 GGCGGCGCCG GTGTGGGCAA GACCGTGATC ATCATGGAGC TGATTAACAA 450
CGTCGCGAAG GGCCACGGTG GTTTCTCGGT GTTTGCCGGC GTTGCGGAGC 500
GCACGCGCGA GGGCACGGAC CTGTACCTGG AGATGATGCA GTCGAAGGTG 550
ATTGACCTGA AGGGCGAGTC GAAGTGCGTG CTTGTGTACG GGCAGATGAA 600
CGAGCCCCCG GGTGCGCGCG CGCGCGTTGC GCAGTCTGCG CTGACGATGG 650
15 CCGAGTACTT CCGAGACGTG GAGGGCCAGA ATGTGCTGCT GTTCATCGAC 700
AACATCTTCC GCTTCACGCA GCGCAACTCC GAGGTCTCTG CGCTGCTGGG 750
CCGCATTCCG CCGCGCCGTG GCTACCAGCC GACGCTTGCG GAGGATCTTG 800
GTATGCTGCA GGAGCGCATC ACGTCGACGA CGAAGGGGTC GATCACGTCC 850
GTGCAGGCCG TGTACGTGCC TGC GGATGAT ATCACGGATC CGGCGCCCCG 900
20 GACGACGTTT TCGCACCTGG ACGCGACGAC TGTGCTGGAC CGCGCGGTGG 950
CGGAGTCGGG GATCTACCC T GCGTGAACC CGCTGGAGTG CGCGTCGCGT 1000
ATCATGGACC CCGATGTGAT CGACGTGGAC CACTACAACG TTGCGCAGGA 1050
TATCGTGCAG ATGCTGACCA AGTACAAGGA GCTGCAGGAT ATCATTGCGG 1100
TGCTTGGTAT CGACGAGCTG AGCGAGGAGG ACAAGGTCGT GGTGGACCGC 1150
25 GCGCGCAAGG TGACCCGGTT CCTGTCGCAG CCGTTCAGG TTGCGGAGGT 1200
GTTCACGGGC ATGACGGGCC AC 1222

```

30 2) INFORMATION FOR SEQ ID NO: 861

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Leishmania tropica*  
 (B) STRAIN: ATCC 50129

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 861

```

45 CCGTGCTGAC GCGCTGGAT GTGACGGAGG ACCTTGGCCG CGATGAGCCG 50
CTGACGCTGG AGATCGTGCA GCACTTGGAC GCGAACACCG GCCGCTGCAT 100
TGCGATGCAG ACGACGGACC TGCTGAAGCT GAAGTCGAAG GTTGTGTCGA 150
CCGGCGGCAA CATCTCTGTG CCGGTGGGCC GTGAGACGCT GGGCCGCATC 200
50 TTCAACGTTT TGGGCGACGC GATCGACCAG CGCGGCCCGG TGGGCGAGAA 250
GATGCGCATG GCGATCCACG CCGAGGCCCG GAAGCTGGCG GATCAGGCCG 300
CGGAGGACAC GATCCTGACG ACCGGCATCA AGGTGATCGA CCTGATTCTG 350
CCCTACTGCA AGGGTGGCAA GATCGGCCTG TTCGGCGGTG CCGGTGTGGG 400
CAAGACTGTG ATCATCATGG AGCTGATCAA CAACGTCGCG AAGGGCCACG 450
55 GTGGTTTCTC CGTGTTTGCC GGC GTTGGCG AGCGCACGCG CGAGGGCACG 500
GACCTGTACC TGGAGATGAT GCAGTCGAAG GTGATTGACC TGAAGGGCGA 550
GTCGAAGTGC GTGCTTGTGT ACGGGCAGAT GAACGAGCCC CCGGTGCGC 600
GCGCGCGCGT TCGCAGTCT GCGCTGACGA TGGCGGAGTA CTTCCGCGAC 650
GTGGAGGGCC AGAACGTGCT GCTGTTCATC GACAACATCT TCCGCTTCAC 700
60 GCAGGCGAAC TCCGAGGTGT CTGCGCTGCT GGGCCGCATT CCGGCCGCCG 750

```

	TGGGCTACCA	GCCGACGCTT	GCGGAGGATC	TTGGTATGCT	GCAGGAGCGC	800
	ATCACGTCGA	CAACGAAGGG	GTCGATCACG	TCCGTGCAGG	CCGTGTACGT	850
	GCCAGCGGAT	GATATCACGG	ATCCC GCGCC	CGCGACGACG	TTCTCGCACC	900
	TGGACGCGAC	GACTGTGCTG	GACCGCGCGG	TGGCGGAGTC	GGGCATCTAC	950
5	CCTGCCGTGA	ACCCGCTGGA	GTGCGCGTCG	CGTATCATGG	ACCCCGATGT	1000
	GATCGATGTG	GACCACTACA	ACGTTGCGCA	GGATATCGTG	CAGATGCTGA	1050
	CCAAGTACAA	GGAGCTGCAG	GATATCATTG	CGGTGCTTGG	CATCGACGAG	1100
	CTGAGCGAGG	AAGACAAGGT	TGTTGTGGAC	CGCGCGCGCA	AGGTGACCCG	1150
	GTTCTGTGTCG	CAGCCGTTCC	AGGTTGCGGA	GGTGTTCACG	GGCATGACGG	1200
10	GCCACTACGT	GCAGCTGGTC	GACACGGTGG	AGTCGTTCTC	TGGCCT	1246

## 2) INFORMATION FOR SEQ ID NO: 862

15

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1265 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

20

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

25

- (A) ORGANISM: *Leishmania tropica*  
 (B) STRAIN: ATCC 30815

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 862

30	GGCGTGCCGC	CCGTGCTGAC	GCGGCTGGAT	GTGACGGAGG	ACCTTGGCCG	50
	CGATGAGCCG	CTGACGCTGG	AGATCGTGCA	GCACTTGGAC	GCGAACACGG	100
	GCCGCTGCAT	TGCGATGCAG	ACGACGGACC	TGCTGAAGCT	GAAGTCGAAG	150
	GTCGTGTGCA	CCGGCGGCAA	CATCTCTGTG	CCGGTGGGCC	GTGAGACGCT	200
	GGGCCGCAATC	TTCAAYGTTC	TGGGCGACGC	GATCGACCAG	CGCGGCCCCG	250
35	TGGGCGAGAA	GATGCGCATG	GCGATCCACG	CCGAGGCCCC	GAAGCTGGCG	300
	GATCAGGCCG	CGGAGGACAC	GATCCTGACG	ACCGGCATCA	AGGTGATCGA	350
	CCTGATTCTG	CCCTACTGCA	AGGGTGGCAA	GATCGGCCTG	TTCGGCGGTG	400
	CCGGTGTGGG	CAAGACTGTG	ATCATCATGG	AGCTGATCAA	CAACGTCGCG	450
	AAGGGCCACG	GCGGTTTCTC	CGTGTTTGCC	GGCGTTGGCG	AGCGCACGCG	500
40	CGAGGGCACG	GACCTGTACC	TGGAGATGAT	GCAGTCGAAG	GTGATTGACC	550
	TGAAGGGCGA	GTCGAAGTGY	GTGCTTGTGT	ATGGGCAGAT	GAACGAGCCC	600
	CCGGGTGCGC	GCGCGCGCGT	TGCGCAGTCT	GCGCTGACGA	TGGCGGAGTA	650
	CTTCCGCGAC	GTGGAGGGCC	AGAACGTGCT	GCTGTTTCATC	GACAACATCT	700
	TCCGCTTCAC	GCAGGCGAAC	TCCGAGGTGT	CTGCGCTGCT	GGGCCGCATT	750
45	CCGGCCGCCG	TGGGCTACCA	GCCGACGCTT	GCGGAGGATC	TTGGTATGCT	800
	GCAGGAGCGC	ATCACGTCGA	CAACGAAGGG	GTCGATCACG	TCCGTGCAGG	850
	CCGTGTACGT	GCCAGCGGAT	GATATCACGG	ATCCC GCGCC	CGCGACGACG	900
	TTCTCGCACC	TGGACGCGAC	GACTGTGCTG	GACCGCGCGG	TGGCGGAGTC	950
	GGGCATCTAC	CCTGCCGTGA	ACCCGCTGGA	GTGCGCGTCG	CGTATCATGG	1000
50	ACCCTGATGT	GATCGATGTG	GACCACTACA	ACGTTGCGCA	GGATATCGTG	1050
	CAGATGCTGA	CCAAGTACAA	GGAGCTGCAG	GATATCATTG	CGGTGCTTGG	1100
	CATCGACGAG	CTGAGCGAGG	AAGACAAGGT	TGTTGTGGAC	CGCGCGCGCA	1150
	AGGTGACCCG	GTTCTGTGTCG	CAGCCGTTCC	AGGTTGCGGA	GGTGTTCACG	1200
	GGCATGACGG	GCCACTACGT	GCAGCTGGTC	GACACGGTGG	AGTCGTTCTC	1250
55	TGGCCTGCTG	ATGGG				1265

## 2) INFORMATION FOR SEQ ID NO: 863

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1191 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bordetella pertussis*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 863

```

15 ATGGCAAAAG GCAAGTTTGA ACGTACCAAG CCGCACGTGA ACGTGGGTAC      50
   GATTGGTCAC GTTGACCACG GCAAAACGAC GTTGACGGCG GCGATCACGA      100
   CCGGTGCTGTC GAACAAGTTC GGCGGCGAGG CTCGCGGCTA CGACCAGATT      150
   GACGCGGCGC CGGAAGAGAA GGCGCGTGGG ATCACGATCA ACACCTCGCA      200
   CGTTGAGTAC GAGACGGAGA CGCGTCACTA CGCGCACGTT GATTGCCCGG      250
   GTCACGCTGA CTACGTGAAG AACATGATCA CGGGTGCTGC GCAGATGGAC      300
20 GGCGCGATCC TGGTGGTGTC GGCCGCAGAC GGCCCGATGC CGCAGACGCG      350
   CGAGCACATT TTGCTGTCGC GCCAGGTTGG CGTGCCGTAC ATCATCGTGT      400
   TCCTGAACAA GGCGGACATG GTTGATGACG CGGAGCTGCT CGAGCTGGTG      450
   GAGATGGAAG TCCGCGAACT GCTGAGCAAG TACGATTTCC CGGGCGATGA      500
   CACGCCGATC GTGAATGGTG CGGCCAAGCT GGCGCTGGAA AGCGACAACG      550
25 GCGACCTGGG CGAGCAGGCG ATTCTGTCGC TGGCGCAAGC GCTGGACACG      600
   TACATTCCGA CGCCGGAGCG CGCGGTCGAC GGTGCGTTCC TGATGCCGGT      650
   GGAAGACGTG TTCTCGATCT CGGGCCGTGG CACGGTGGTG ACTGGCCGTA      700
   TCGAGCGCGG CGTGGTGAAG GTTGCGCAGG AAATCGAAAT CGTGGGCATC      750
   AAGCCGACGG TGAAGACGAC CTGCACGGGC GTGGAGATGT TCCGCAAGCT      800
30 GCTGGACCAG GGCCAGGCGG GCGACAACGT GGGTATCTTG CTGCGCGGCA      850
   CCAAGCGTGA AGACGTCGAG CGTGGCCAGG TGCTGGCCAA GCCGGGTTTC      900
   ATCAACCCGC ACACGGACTT CACGGCCGAG GTGTACATTC TGTCCAAGGA      950
   AGAGGGTGGC CGTCACACGC CGTTCTTCAA CGGCTATCGT CCGCAGTTCT     1000
   ACTTCCGCAC GACGGACGTG ACCGGCACGA TCGACCTGCC GGCGGACAAG     1050
35 GAAATGGTGC TGCCGGGCGA CAACGTGTCT ATGACCGTCA AGCTGCTGGC     1100
   CCCGATCGCC ATGGAAGAAG GTCTGCGTTT CGCCATCCGT GAAGGCGGTC     1150
   GTACCGTCGG TGCCGGCGTC GTCGCCAAGA TCATCAAGTA A              1191

```

## 2) INFORMATION FOR SEQ ID NO: 864

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1350 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Trypanosoma brucei*  
 (B) STRAIN: LVH/75/USAMRU-K/18  
 (C) ACCESSION NUMBER: extracted from U10562

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 864

```

60 ATGGGAAAGG AAAAGGTGCA CATGAATCTT GTGGTGGTGG GCCACGTGCA      50
   TGCCGGTAAA TCCACTGCAA CGGGTCACTT GATCTACAAG TGCGGTGGTA      100
   TTGACAAACG TACGATCGAG AAGTTCGAGA AAGAAGCTGC CGACATTGGT      150

```

	AAGGCCTCAT	TCAAGTACGC	ATGGGTGCTG	GACAAGCTGA	AGGCTGAGCG	200
	CGAACGTGGT	ATCACGATCG	ACATTGCACT	GTGGAAATTC	GAGTCACCCA	250
	AGTCTGTCTT	CACTATTATT	GATGCTCCTG	GGCACCGTGA	CTTCATCAAG	300
	AACATGATCA	CCGGCACATC	GCAAGCCGAC	GCAGCCATCC	TCATCATTGC	350
5	CTCTGCGCAG	GGTGAGTTTC	AGGCTGGTAT	CTCCAAGGAT	GGACAGACCC	400
	GCGAGCACGC	GTTGCTGGCC	TTCACTTTGG	GTGTGAAGCA	GATGGTTGTG	450
	TGCTGCAACA	AAATGGACGA	CAAGACTGTG	AACTACGGAC	AGGAGCGGTA	500
	TGACGAGATT	GTGAAGGAGG	TGTCTGCTTA	CATCAAGAAG	GTTGGGTACA	550
	ACGTGGAGAA	GGTGCGCTTC	GTCCCCATCT	CCGGATGGCA	GGGCGACAAC	600
10	ATGATTGAGA	AATCCGAGAA	GATGCCATGG	TACAAGGGTC	CAACGCTCCT	650
	GGAGGCACTA	GACATGCTGG	AGCCACCAGT	GCGTCCGAGC	GACAAGCCCC	700
	TGCGTCTGCC	ACTGCAGACG	TGTACAAAGA	TCGGTGGTAT	TGGCACCCTG	750
	CCCGTTGGTC	GTGTGGAGAC	CGGCGTGATG	AAGCCTGGTG	ATGTGGTGAC	800
	GTTTGCCCCC	GCCAACGTGA	CGACCGAGGT	GAAATCGATC	GAGATGCACC	850
15	ACGAGCAGCT	CGCTGAGGCG	ACCCCCGGTG	ACAACGTCGG	CTTTAACGTG	900
	AAGAACGTTT	CTGTAAAGGA	CATCCGCCGT	GGCAACGTCT	GCGGTAACAC	950
	CAAGAACGAC	CCCCCAAAGG	AGGCCGCCGA	CTTCACGGCA	CAGGTGATCA	1000
	TCCTGAACCA	CCCCGGACAG	ATTGGAAACG	GTTATGCGCC	CGTGCTGGAC	1050
	TGCCACACAT	CGCACATTGC	CTGCAAGTTC	GCGGAGATCG	AGTCGAAGAT	1100
20	CGACCGTCGC	TCTGGCAAGG	AGCTGGAGAA	GGCTCCCAAG	TCGATCAAGT	1150
	CTGGCGACGC	CGCGATCGTG	CGCATGGTGC	GCGAGAAGCC	TATGTGCGTG	1200
	GAGGTC'TTCA	ACGACTACGC	GCCACTCGGC	CGCTTTGCCG	TGCGTGACAT	1250
	GCGCCAGACC	GTCGCTGTCT	GTATCATCAA	GGCCGTGACC	AAGAAGGACG	1300
25	GTTCTGGTGG	TAAGGTGACG	AAGGCTGCGG	TGAAGGCTTC	GAAGAAATAA	1350

## 2) INFORMATION FOR SEQ ID NO: 865

## 30 (i) SEQUENCE CHARACTERISTICS:

(A) —LENGTH: 1052—bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Cryptosporidium parvum*

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 865

	AAGCTCAAGG	CTGAGAGAGA	AAGAGGTATT	ACCATCGATA	TTGCTTTTATG	50
	GCAATTCGAA	ACCCCAAAAT	ACCACTACAC	TGTCATTGAT	GCCCCAGGTC	100
45	ACAGAGATTT	CATCAAGAAT	ATGATTACTG	GTACCTCTCA	AGCTGATGTT	150
	GCTTTTATGG	TTGTCCCAGC	CGATCGTTTC	GAAGGTGCCT	TCTCCAAGGA	200
	AGGTCAAACC	AGAGAACATG	CTTTATTGGC	CTTCACTTTG	GGTGTCAGAC	250
	AAATGATTGT	CGGTATTAAC	AAGATGGATA	CCTGTGAATA	CAAGCAATCT	300
	CGTTTTTGATG	AAATCTTCAA	CGAAGTTGAT	GGTTACCTCA	AGAAGGTTGG	350
50	TTACAACACC	GAGAAGATCC	CATTCTGTTG	CATTTCTGGT	TTCTGTTGGTG	400
	ATAATATGGT	TGAGAGATCT	GACAAGATGC	CATGGTATAA	GGGTAAGACC	450
	TTAGTCGAAG	CCCTCGACAC	TATGGAACCA	CCAAAGAGAC	CAACTGACAA	500
	GCCACTCCGT	CTCCCATTAC	AAGATGTTTA	CAAGATAGGT	GGTGTAAGTA	550
	CTGTCCCAGT	CGGTCGTGTT	GAGACTGGTA	TCATCAGACC	AGGTATGAAT	600
55	GTTACCTTCG	CTCCAGCTGG	TGTTACCACT	GAAGTTAAGT	CAGTAGAAAT	650
	GCACCATGAG	CAGATGCCAG	AGGCCGTCCC	AGGTGACAAC	GTTGGTTTCA	700
	ATGTTAAGAA	CGTCTCCATC	AAGGATATCA	AGAGAGGTTT	CGTTGCTTCT	750
	GATGCCAAGA	ATGACCCAGC	TAAGGGCTGT	GAAGACTTCA	CTGCTCAAGT	800
	TATCGTCCCT	AACCACCCAG	GTGAAATCAA	GAACGGTTAC	TCTCCAGTCG	850
60	TTGACTGTCA	CACCGCTCAC	ATTCCTTGCA	AATTCCAGAC	TATCACTGCT	900



AAGATGGACA	AGAGATCTGG	TAAGGTTTTG	GAAGAAAACC	CAAAGCTTAT	950
CAAGTCTGGT	GATGCTGCTT	TGGTTGTTAT	GCAACCTTTG	AAGCCACTTT	1000
GTGTTGAGGC	CTTCACTGAC	TACCCACCTC	TAGGTCGTTT	CGCTGTCCGT	1050
GA					1052

5

## 2) INFORMATION FOR SEQ ID NO: 866

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 837 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: *Staphylococcus saprophyticus*  
 (B) STRAIN: ATCC 35552

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 866

25	CAATGAAGTT	CCAGAAATTA	ACAATGCCTT	AGTCGTAGAC	GTTGAAAGAG	50
	ATGAAGGTAC	AGTATCTCTT	ACATTAGAAG	TGGCATTACA	ACTTGGCGAT	100
	GATGTCGTAC	GTACAATTGC	AATGGATTCT	ACTGATGGTG	TTAAACGTGG	150
	TACAGAAAGTT	CGAGATAGCG	GAGATAGCAT	CAGTGTTCCT	GTTGGTGTATG	200
	CTACGTTAGG	ACGTGTGTTT	AATGTTCTTG	GTGATACAAT	TGACTTAGAC	250
	GAGAAGCTTG	ATACTTCTGT	CAAACGTGAT	CCAATTCATA	GAGAAGCACC	300
30	TGCATTGAT	CAATTATCAA	CAAAAGTTGA	AATCTTAGAA	ACAGGTATTA	350
	AAGTAATTGA	TTTACTTGCA	CCATATATTA	AAGGTGGTAA	AATCGGTTTA	400
	TTCGGTGGCG	CTGGTGTAGG	TAAAACAGTA	TTAATTCAAG	AATTAATTAA	450
	TAATATAGCT	CAAGAACATG	GTGGTATTTT	AGTATTTGCC	GGCGTAGGTG	500
	AACGTACGCG	TGAAGGTAAT	GACTTATACT	ACGAAATGAG	TGATAGTGGT	550
35	GTTATTAAGA	AAACAGCTAT	GGTCTTCGGA	CAAATGAATG	AGCCACCTGG	600
	TGCGCGTATG	CGTGTGCTT	TATCAGGCTT	AACAATGGCT	GAACACTTCC	650
	GTGATGTACA	AGGACAAGAT	GTTTTACTAT	TTATTGATAA	CATATTCAGA	700
	TTTACGCAAG	CTGGTTCAGA	AGTATCAGCA	CTATTAGGTC	GTATGCCATC	750
	AGCCGTTGGT	TATCAACCTA	CCCTTGCTAC	TGAAATGGGT	CAATTACAAG	800
40	AACGTATTAC	ATCAACAAC	AAAGGATCTG	TAACGTC		837

## 2) INFORMATION FOR SEQ ID NO: 867

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 818 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 50 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

55 (A) ORGANISM: *Zoogloea ramigera*  
 (B) STRAIN: ATCC 25935

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 867

60	AAGGTATTCG	ATGCCTTGAA	AATGGAAGGC	TCCGAGCTGA	CCCTGGAAGT	50
----	------------	------------	------------	------------	------------	----

690

```

5   ACAACAGCAG CTGGGCGACG GCATTGTCCG TACCATTGCA CTGGGTACCT 100
    CCGACGGCCT GCGTCGCGGC ATGATGATCC AGAACACCGG CAAACCTATC 150
    ATGGTGCCAG TCGGTAAAGC AACCTGGGGT CGCATCATGG ACGTGCTGGG 200
    TAACCCGATC GACGAATGCG GCGCGGTCGC TCACGACCAG ATCGCTTCGA 250
    TCCACCGCGC TCCTCCTGCG TACGACGAAC TGTCGCCATC GCAAGATCTG 300
    CTGGAAACCG GCATTAAAGT TATTGACCTG GTGTGCCCGT TCGCCAAGGG 350
    CGGTAAAGTC GGTCTGTTCG GCGGTGCAGG TGTGGGCAAG ACCGTGAACA 400
    TGATGGAAC TATCAACAAC ATCGCCAAAG CACACTCGGG TCTGTCCGTG 450
    TTTGCCGGTG TGGGTGAGCG TACCCGTGAA GGTAACGACT TCTACCACGA 500
10  GATGGCTGAC GCCAAAGTGG TCGATCTGGA AAATCCAGAG AACTCCAAGG 550
    TTGCGATGGT CTACGGTCAG ATGAATGAAC CACCAGGCAA CCGTCTGCGC 600
    GTGGCGCTGA CCGGTCTGAC CATGGCTGAA GCATTCCGTG ACGAAGGCAA 650
    AGACGTTCTG TTCTTCGTGG ACAACATCTA CCGCTTCACC CTGGCCGGTA 700
    CCGAAGTATC GGCACGTCTG GGCCGTATGC CATCGGCTGT GGGTTACCAG 750
15  CCTACGCTGG CCGAAGAAAT GGGTCGCCTG CAAGAGCGCA TCACTTCGAC 800
    CAAGACCGGT TCGATCAC 818

```

## 20 2) INFORMATION FOR SEQ ID NO: 868

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 778 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus saprophyticus*  
 (B) STRAIN: ATCC 43867

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 868

```

35  CTATCTTAGT AGTATCTGCT GCTGATGGCC CAATGCCACA AACTCGTGAA 50
    CACATTCTTT TATCACGTAA CGTTGGTGTT CCAGCATTAG TTGTATTCTT 100
    AAACAAAGTT GACATGGTTG ACGATGAAGA ATTATTAGAA TTAGTAGAAA 150
    TGGAAGTTTCG TGACTTATTA AGCGAATATG ACTTCCCAGG TGACGATGTA 200
40  CCTGTAATCT CTGGTTCTGC ATTAAAAGCT TTAGAAGGCG ACGCTGACTA 250
    TGAGCAAAAA ATCTTAGACT TAATGCAAGC TGTGATGAC TTCATTCCAA 300
    CACCAGAACG TGATTCTGAC AAACCATTCA TGATGCCAGT TGAGGACGTA 350
    TTCTCAATCA CTGGTCGTGG TACTGTTGCT ACAGGCCGTG TTGAACGTGG 400
    TCAAATCAAA GTCGGTGAAG AAATCGAAAT CATCGGTATG CAAGAAGAAT 450
45  CAAGCAAAAC AACTGTTACT GGTGTAGAAA TGTTCGTAA ATTATTAGAC 500
    TACGCTGAAG CTGGTGACAA CATTGGTGCA TTATTACGTG GTGTTTCACG 550
    TGATGACGTA CAACGTGGTC AAGTTTTAGC TGCTCCTGGT ACTATTACAC 600
    CACATACAAA ATTCAAAGCG GATGTTTACG TTTTATCTAA AGATGAAGGT 650
    GGTCGTCATA CACATTCTT CACTAATAC CGCCACAAT TCTATTTCCG 700
50  TACTACTGAC GTAACGTGGT TTGTTAACTT ACCAGAAGGT ACTGAAATGG 750
    TTATGCCTGG CGATAACGTT GAAATGGA 778

```

## 55 2) INFORMATION FOR SEQ ID NO: 869

## (i) SEQUENCE CHARACTERISTICS:

- 60 (A) LENGTH: 640 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus casseliflavus*  
(B) STRAIN: R689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 869

```

10 TGGTCCTATG CCTCAAACAC GTGAACACAT CTTGTTATCA CGTAACGTTG      50
   GTGTACCATA CATCGTTGTT TTCTTAAACA AAATGGATAT GGTTGATGAC      100
   GAAGAATTAC TAGAATTAGT TGAAATGGAA GTTCGTGACT TATTGTCAGA      150
   ATATGACTTC CCAGGCGACG ATGTTCTGT AATCGCTGGT TCTGCTTTGA      200
15 AAGCTCTTGA AGGCGATGCT TCATACGAAG AAAAAATCAT GGAATTAATG      250
   GCTGCAGTTG ACGAATACGT TCCAACCTCA GAACGTGACA CTGACAAACC      300
   ATTCATGATG CCAGTCGAAG ACGTATTCTC AATCACTGGA CGTGGTACTG      350
   TTGCTACAGG CCGTGTGAA CGTGGACAAG TTCGCGTTGG TGACGAAGTT      400
   GAAATCGTTG GTATTGCTGA AGAACTGCT AAAACAACCTG TAACTGGTGT      450
20 TGAAATGTTT CGTAAATTGT TAGACTATGC TGAAGCAGGG GATAACATTG      500
   GTGCATTGCT ACGTGGTGT GTCGTGAAG ACATCCAACG TGGACAAGTA      550
   TTGGCTAAAG CTGGTACAAT CACACCTCAT ACAAATTTTA AAGCTGAAGT      600
   TTACGTTTTA ACAAAGAAG AAGGTGGACG TCACACTCCA      640

```

25

2) INFORMATION FOR SEQ ID NO: 870

(i) SEQUENCE CHARACTERISTICS:

```

30 (A) LENGTH: 644 bases
   (B) TYPE: Nucleic acid
   (C) STRANDEDNESS: Double
   (D) TOPOLOGY: Linear

```

35 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus casseliflavus*  
(B) STRAIN: R754

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 870

```

   GTCCTATGCC TCAAACACGT GAACACATCT TGTTATCACG TAACGTTGGT      50
   GTACCATACTA TCGTTGTTTT CTAAACAAA ATGGATATGG TTGATGACGA      100
45 AGAATTACTA GAATTAGTTG AAATGGAAGT TCGTGAAGTA TTGTCAGAAT      150
   ATGACTTCCC AGGCGACGAT GTTCCTGTAA TCGCTGGTTC TGCTTTGAAA      200
   GCTCTTGAAG GCGATGCTTC ATACGAAGAA AAAATCATGG AATTAATGGC      250
   TGCAGTTGAC GAATACGTTT CAACTCCAGA ACGTGACACT GACAAACCAT      300
   TCATGATGCC AGTCGAAGAC GTATTCTCAA TCACTGGACG TGGTACTGTT      350
50 GCTACAGGCC GTGTTGAACG TGGACAAGTT CGCGTTGGTG ACGAAGTTGA      400
   AATCGTTGGT ATTGCTGAAG AACTGCTAA AACAACGTGA ACTGGTGTG      450
   AAATGTTCCG TAAATTGTTA GACTATGCTG AAGCAGGGGA TAACATTGGT      500
   GCATTGCTAC GTGGTGTGTC TCGTGAAGAC ATCCAACGTG GACAAGTATT      550
   GGCTAAAGCT GGTACAATCA CACCTCATAC AAAATTTAAA GCTGAAGTTT      600
55 ACGTTTTAAC AAAAGAAGAA GGTGGACGTC ACACACCATT CTTC      644

```

60

2) INFORMATION FOR SEQ ID NO: 871

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 637 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus flavescens*  
 (B) STRAIN: R758

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 871

```

15 TCCTATGCCT CAAACACGTG AACACATCTT GTTATCACGT AACGTTGGTG      50
   TACCATACAT CGTTGTTTTT TTAACAACAA TGGATATGGT TGATGACGAA      100
   GAATTACTAG AATTAGTTGA AATGGAAGTT CGTGACTTAT TGTCAGAATA      150
   TGAATTCCCA GGCGACGATG TTCCTGTAAT CGCTGGTTCT GCTTTGAAAG      200
   CTCTTGAAGG CGATGCTTCA TACGAAGAAA AAATCATGGA ATTAATGGCT      250
20 GCAGTTGACG AATACGTTCC AACTCCAGAA CGTGACACTG ACAAACCATT      300
   CATGATGCCA GTCGAAGACG TATTCTCAAT CACTGGACGT GGTACTGTTG      350
   CTACAGGCCG TGTTGAACGT GGACAAGTTC GCGTTGGTGA CGAAGTTGAA      400
   ATCGTTGGTA TTGCTGAAGA AACTGCTAAA ACAACTGTAA CTGGTGTTGA      450
   AATGTTCCGT AAATTGTTAG ACTATGCTGA AGCAGGGGAT AACATTGGTG      500
25 CATTGCTACG TGGGGTTGCT CGTGAAGACA TCCAACGTGG ACAAGTATTA      550
   GCTAAAGCTG GTACAATCAC ACCTCATACA AAATTTAAAG CTGAAGTTTA      600
   CGTTTTAACA AAAGAAGAAG GTGGACGTCA CACTCCA      637

```

## 2) INFORMATION FOR SEQ ID NO: 872

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 643 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus gallinarum*  
 (B) STRAIN: R631

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 872

```

GTCCTATGCC TCAAACCTCGT GAACACATCT TGTTATCACG TAACGTTGGC      50
GTACCATACA TCGTTGTTTT CTTGAACAAA ATGGATATGG TTGATGACGA      100
AGAATTGCTA GAATTAGTTG AAATGGAAGT TCGTGACCTA TTGTCTGAAT      150
50 ATGACTTCCC AGGCGACGAT GTTCTGTAA TCGCCGGTTC TGCTTTGAAA      200
   GCTCTTGAAG GAGATCCTTC ATACGAAGAA AAAATCATGG AATTGATGGC      250
   TGCAGTTGAC GAATACGTTT CAACTCCAGA ACGTGATACT GACAAACCAT      300
   TCATGATGCC AGTCGAAGAC GTATTCTCAA TCACTGGACG TGGTACTGTT      350
   GCTACAGGCC GTGTTGAACG TGGACAAGTT CGCGTTGGTG ATGAAGTAGA      400
55 AATCGTTGGT ATTGCTGACG AACTGCTAA AACAACGTGA ACAGGTGTTG      450
   AAATGTTCCG TAAATTGTTA GACTATGCTG AAGCAGGGGA TAACATTGGT      500
   GCATTGCTAC GTGGGGTTGC TCGTGAAGAC ATCCAACGTG GACAAGTATT      550
   GGCTAAAGCT GGTACAATCA CACCTCATAC AAAATTCAAA GCTGAAGTTT      600
   ATGTTTTGAC AAAAGAAGAA GGTGGACGTC AACTCCATT CTT      643

```

## 2) INFORMATION FOR SEQ ID NO: 873

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 641 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 10 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Enterococcus gallinarum*  
 15 (B) STRAIN: R691

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 873

	GTCCTATGCC	TCAAACCTCGT	GAACACATCT	TGTTATCACG	TAACGTTGGC	50
20	GTACCATACA	TCGTTGTTTT	CTTGAACAAA	ATGGATATGG	TTGATGACGA	100
	AGAATTGCTA	GAATTAGTTG	AAATGGAAGT	TCGTGACCTA	TTGTCTGAAT	150
	ATGACTTCCC	AGGCGACGAT	GTTCTGTAA	TCGCCGGTTC	TGCTTTGAAA	200
	GCTCTTGAAG	GAGATCCTTC	ATACGAAGAA	AAAATCATGG	AATTGATGGC	250
	TGCAGTTGAC	GAATACGTTT	CAACTCCAGA	ACGTGATACT	GACAAACCAT	300
25	TCATGATGCC	AGTCGAAGAC	GTATTCTCAA	TCACTGGACG	TGGTACTGTT	350
	GCTACAGGCC	GTGTTGAACG	TGGACAAGTT	CGCGTTGGTG	ATGAAGTAGA	400
	AATCGTTGGT	ATTGCTGACG	AAACTGCTAA	AACAACGTGA	ACAGGTGTTG	450
	AAATGTTCCG	TAAATTGTTA	GACTATGCTG	AAGCAGGGGA	TAACATTGGT	500
	GCATTGCTAC	GTGGGGTTGC	TCGTGAAGAC	ATCCAACGTG	GACAAGTATT	550
30	GGCTAAAGCT	GGTACAATCA	CACCTCATAC	AAAATTCAAA	GCTGAAGTTT	600
	ATGTTTTGAC	AAAAGAAGAA	GGTGGACGTC	ACACTCCATT	C	641

## 35 2) INFORMATION FOR SEQ ID NO: 874

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 681 bases  
 (B) TYPE: Nucleic acid  
 40 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- 45 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Staphylococcus haemolyticus*  
 (B) STRAIN: LSPQ 2514

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 874

50	ACCAGCATT	A	GTAGTATTCT	TAAATAAAGT	TGACATGGTT	GACGATGAAG	50
	AATTATTAGA	ATTAGTTGAA	ATGGAAGTAC	GTGACTTATT	ATCTGAATAC		100
	GACTTCCCAG	GTGACGATGT	ACCTGTAATC	GCTGGTTCAG	CATTAAAAGC		150
	TTTAGAAGGC	GATGCTCAAT	ACGAAGAAAA	AATCTTAGAA	TTAATGCAAG		200
55	CAGTTGATGA	CTACATTCCA	ACTCCAGAAC	GTGATTCTGA	CAAACCATTC		250
	ATGATGCCAG	TTGAGGACGT	ATTCTCAATC	ACTGGTCGTG	GTACTGTTGC		300
	TACAGGCCGT	GTTGAACGTG	GGCAAATCAA	AGTTGGTGAA	GAAGTTGAAA		350
	TCATTGGTAT	CCATGACACT	TCTAAAACAA	CTGTTACTGG	TGTAGAAATG		400
	TTCCGTAAAT	TATTAGACTA	CGCTGAAGCT	GGTGACAACA	TCGGTGCATT		450
60	ATTACGTGGT	GTTGCTCGTG	AAGACGTACA	ACGTGGTCAA	GTATTAGCTG		500

CTCCAGGTTC	AATCACACCT	CACACAAAAT	TTAAAGCAGA	CGTATACGTT	550
TTATCTAAAG	ACGAAGGTGG	ACGTCACACT	CCATTCTTCA	CAAACATATCG	600
TCCACAATTC	TATTTCCGTA	CTACTGACGT	AACTGGTGTG	GTTAACCTAC	650
CAGAAGGTAC	TGAAATGGTT	ATGCCTGGCG	A		681

5

## 2) INFORMATION FOR SEQ ID NO: 875

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 675 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM: *Staphylococcus epidermidis*  
 (B) STRAIN: R591

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 875

ATTATCACGT	AACGTTGGTG	TACCAGCATT	AGTTGTATTC	TTAAACAAAG	50
TTGACATGGT	AGACGACGAA	GAATTATTAG	AATTAGTTGA	AATGGAAGTT	100
CGTGACTTAT	TAAGCGAATA	TGACTTCCCA	GGTGACGATG	TACCTGTAAT	150
CGCTGGTTCT	GCATTAAAAG	CATTAGAAGG	CGATGCTGAA	TACGAACAAA	200
AAATCTTAGA	CTTAATGCAA	GCAGTTGATG	ATTACATTCC	AACTCCAGAA	250
CGTGATTCTG	ACAAACCATT	CATGATGCCA	GTTGAGGACG	TATTCTCAAT	300
30 CACTGGTCGT	GGTACTGTTG	CTACAGGCCG	TGTTGAACGT	GGTCAAATCA	350
AAGTTGGTGA	AGAAGTTGAA	ATCATCGGTA	TGCACGAAAC	TTCTAAAACA	400
ACTGTTACTG	GTGTAGAAAT	GTTCCGTAAA	TTATTAGACT	ACGCTGAAGC	450
TGGTGACAAC	ATCGGTGCTT	TATTACGTGG	TGTTGCACGT	GAAGACGTAC	500
AACGTGGTCA	AGTATTAGCT	GCTCCTGGTT	CTATTACACC	ACACACAAAA	550
35 TTCAAAGCTG	AAGTATACGT	ATTATCTAAA	GATGAAGGTG	GACGTCACAC	600
TCCATTCTTC	ACTAACTATC	GCCCACAATT	CTATTTCCGT	ACTACTGACG	650
TAAGTGGTGT	TGTAAACTTA	CCAGA			675

40

## 2) INFORMATION FOR SEQ ID NO: 876

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 704 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Staphylococcus epidermidis*  
 (B) STRAIN: CSG 10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 876

TCTTATTATC	ACGTAACGTT	GGTGTACCAG	CATTAGTTGT	ATTCTTAAAC	50
AAAGTTGACA	TGGTAGACGA	CGAAGAATTA	TTAGAATTAG	TTGAAATGGA	100
AGTTCGTGAC	TTATTAAGCG	AATATGACTT	CCCAGGTGAC	GATGTACCTG	150
60 TAATCGCTGG	TTCTGCATTA	AAAGCATTAG	AAGGCGATGC	TGAATACGAA	200

	CAAAAAATCT	TAGACTTAAT	GCAAGCAGTT	GATGATTACA	TTCCAACCTCC	250
	AGAACGTGAT	TCTGACAAAC	CATTCATGAT	GCCAGTTGAG	GACGTATTCT	300
	CAATCACTGG	TCGTGGTACT	GTTGCTACAG	GCCGTGTTGA	ACGTGGTCAA	350
	ATCAAAGTTG	GTGAAGAAGT	TGAAATCATC	GGTATGCACG	AAACTTCTAA	400
5	AACAACCTGTT	ACTGGTGTAG	AAATGTTCCG	TAAATTATTA	GACTACGCTG	450
	AAGCTGGTGA	CAACATCGGT	GCTTTATTAC	GTGGTGTTCG	ACGTGAAGAC	500
	GTACAACGTG	GTCAAGTATT	AGCTGCTCCT	GGTTCTATTA	CACCACACAC	550
	AAAATTCAAA	GCTGAAGTAT	ACGTATTATC	TAAAGATGAA	GGTGGACGTC	600
	ACACTCCATT	CTTCACTAAC	TATCGCCCAC	AATTCTATTT	CCGTACTACT	650
10	GACGTAACCTG	GTGTTGTAAA	CTTACCAGAA	GGTACAGAAA	TGGTTATGCC	700
	TGGC					704

15 2) INFORMATION FOR SEQ ID NO: 877

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 770 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

25 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus epidermidis*  
 (B) STRAIN: ATCC 35984

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 877

30	TCTTAGTTGT	ATCTGCTGCT	GACGGTCCAA	TGCCACAAAC	TCGTGAACAC	50
	ATCTTATTAT	CACGTAACGT	TGGTGTACCA	GCATTAGTTG	TATTCTTAAA	100
	CAAAGTTGAC	ATGGTAGACG	ACGAAGAATT	ATTAGAATTA	GTTGAAATGG	150
	AAGTTCGTGA	CTTATTAAGC	GAATATGACT	TCCCAGGTGA	CGATGTACCT	200
35	GTAATCGCTG	GTTCTGCATT	AAAAGCATTA	GAAGGCGATG	CTGAATACGA	250
	ACAAAAAATC	TTAGACTTAA	TGCAAGCAGT	TGATGATTAC	ATTCCAACCTC	300
	CAGAACGTGA	TTCTGACAAA	CCATTCATGA	TGCCAGTTGA	GGACGTATTC	350
	TCAATCACTG	GTCGTGGTAC	TGTTGCTACA	GGCCGTGTTG	AACGTGGTCA	400
	AATCAAAGTT	GGTGAAGAAG	TTGAAATCAT	CGGTATGCAC	GAAACTTCTA	450
40	AAACAACGTG	TACTGGTGTA	GAAATGTTCC	GTAAATTATT	AGACTACGCT	500
	GAAGCTGGTG	ACAACATCGG	TGCTTTATTA	CGTGGTGTTC	CACGTGAAGA	550
	CGTACAACGT	GGTCAAGTAT	TAGCTGCTCC	TGGTTCTATT	ACACCACACA	600
	CAAAATTCAA	AGCTGAAGTA	TACGTATTAT	CTAAAGATGA	AGGTGGACGT	650
	CACACTCCAT	TCTTCACTAA	CTATCGCCCA	CAATTCTATT	TCCGTACTAC	700
45	TGACGTAAC	GGTGTGTAA	ACTTACCAGA	AGGTACAGAA	ATGGTTATGC	750
	CTGGCGACAA	CGTTGAAATG				770

50 2) INFORMATION FOR SEQ ID NO: 878

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 716 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

60 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Staphylococcus epidermidis*  
 (B) STRAIN: ATCC 35983

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 878

```

5      TTGTATTCTT AAACAAAGTT GACATGGTAG ACGACGAAGA ATTATTAGAA      50
      TTAGTTGAAA TGGAAGTTCG TGAATTATTA AGCGAATATG ACTTCCCAGG      100
      TGACGATGTA CCTGTAATCG CTGGTTCTGC ATTAAAAGCA TTAGAAGGCG      150
      ATGCTGAATA CGAACAAAAA ATCTTAGACT TAATGCAAGC AGTTGATGAT      200
10     TACATTCCAA CTCCAGAACG TGATTCTGAC AAACCATTCA TGATGCCAGT      250
      TGAGGACGTA TTCTCAATCA CTGGTCGTGG TACTGTTGCT ACAGGCCGTG      300
      TTGAACGTGG TCAAATCAAA GTTGGTGAAG AAGTTGAAAT CATCGGTATG      350
      CACGAAACTT CTAAAACAAC TGTTACTGGT GTAGAAATGT TCCGTAAATT      400
      ATTAGACTAC GCTGAAGCTG GTGACAACAT CGGTGCTTTA TTACGTGGTG      450
15     TTGCACGTGA AGACGTACAA CGTGGTCAAG TATTAGCTGC TCCTGGTTCT      500
      ATTACACCAC ACACAAAAAT CAAAGCTGAA GTATACGTAT TATCTAAAGA      550
      TGAAGGTGGA CGTCACACTC CATTCTTCAC TAACTATCGC CCACAATTCT      600
      ATTTCCGTAC TACTGACGTA ACTGGTGTTG TAAACTTACC AGAAGGTACA      650
      GAAATGGTTA TGCCTGGCGA CAACGTTGAA ATGACAGTTG AATTAATCGC      700
20     TCCAATCGCT ATCGAA      716
  
```

## 2) INFORMATION FOR SEQ ID NO: 879

```

25     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 640 bases
          (B) TYPE: Nucleic acid
          (C) STRANDEDNESS: Double
30     (D) TOPOLOGY: Linear

          (ii) MOLECULE TYPE: Genomic DNA

          (vi) ORIGINAL SOURCE:
35     (A) ORGANISM: Enterococcus gallinarum
          (B) STRAIN: R764
  
```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 879

```

40     CGGTCCTATG CCTCAAACCTC GTGAACACAT CTTGTTATCA CGTAACGTTG      50
      GCGTACCATA CATCGTTGTT TTCTTGAACA AAATGGATAT GGTTGATGAC      100
      GAAGAATTGC TAGAATTAGT TGAAATGGAA GTTCGTGACC TATTGTCTGA      150
      ATATGACTTC CCAGGCGACG ATGTTCTGTG AATCGCCGGT TCTGCTTTGA      200
      AAGCTCTTGA AGGAGATCCT TCATACGAAG AAAAAATCAT GGAATTGATG      250
45     GCTGCAGTTG ACGAATACGT TCCAACCTCCA GAACGTGATA CTGACAAACC      300
      ATTCATGATG CCAGTCGAAG ACGTATTCTC AATCACTGGA CGTGGTACTG      350
      TTGCTACAGG CCGTGTTGAA CGTGGACAAG TTCGCGTTGG TGATGAAGTA      400
      GAAATCGTTG GTATTGCTGA CGAAACTGCT AAAACAACCTG TAACAGGTGT      450
      TGAATGTTTC CGTAAATTGT TAGACTATGC TGAAGCAGGG GATAACATTG      500
50     GTGCATTGCT ACGTGGGGTT GCTCGTGAAG ACATCCAACG TGGACAAGTA      550
      TTGGCTAAAG CTGGTACAAT CACACCTCAT ACAAATTCA AAGCTGAAGT      600
      TTATGTTTTG ACAAAGAAG AAGGTGGACG TCACACTCCA      640
  
```

55 2) INFORMATION FOR SEQ ID NO: 880

```

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 831 bases
60     (B) TYPE: Nucleic acid
  
```



(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Pseudomonas aeruginosa*  
(B) STRAIN: PAO-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 880

```

CGGCGCGATC CTGGTTTGCT CGGCTGCCGA CGGCCCCATG CCGCAGACCC      50
GCGAGCACAT CCTGCTGTCC CGCCAGGTAG GCGTTCCTTA CATCGTCGTG      100
TTCCTGAACA AGGCCGACAT GGTCGACGAC GCCGAGCTGC TGGAACTGGT      150
CGAGATGGAA GTTCGCGATC TGCTGAACAC CTACGACTTC CCGGGCGACG      200
ACACTCCGAT CATCATCGGT TCCGCGCTGA TGGCGCTGGA AGGCAAGGAT      250
GACAACGGCA TCGGCGTAAG CGCCGTGCAG AAGCTGGTAG AGACCCTGGA      300
CTCCTACATT CCGGAGCCGG TTCGTGCCAT CGACCAGCCG TTCCTGATGC      350
CGATCGAAGA CGTGTTCTCG ATCTCCGGCC GCGGTACCGT GGTAACCGGT      400
CGTGTAGAGC GCGGCATCAT CAAGGTCCAG GAAGAAGTGG AAATCGTCGG      450
CATCAAGGCG ACCACCAAGA CTACCTGCAC CGGCGTTGAA ATGTTCCGCA      500
AGCTGCTCGA CGAAGGTCGT GCTGGTGAGA ACGTTGGTAT CCTGCTGCGT      550
GGCACCAAGC GTGAAGACGT AGAGCGTGCG CAGGTTCTGG CCAAGCCGGG      600
CACCATCAAG CCGCACACCA AGTTCGAGTG CGAAGTGTAC GTGCTGTCCA      650
AGGAAGAAGG TGGTCGTCAC ACCCCGTTCT TCAAGGGCTA CCGTCCGCAG      700
TTCTACTTCC GTACCACCGA YGTGACCGGT AACTGCGAAC TGCCGGAAGG      750
CGTAGAGATG GTAATGCCGG GCGACAACAT CAAGATGGTT GTCACCCTGA      800
TCGCTCCGAT CGCCATGGAA GATGGCCTGC G          831

```

2) INFORMATION FOR SEQ ID NO: 881

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 642 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus casseliflavus*  
(B) STRAIN: R421

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 881

```

CCTATGCCTC AAACACGTGA ACACATCTTG TTATCACGTA ACGTTGGTGT      50
ACCATACATC GTTGTTTTCT TAAACAAAAT GGATATGGTT GATGACGAAG      100
AATTACTAGA ATTAGTTGAA ATGGAAGTTC GTGACTTATT GTCAGAATAT      150
GACTTCCCAG GCGACGATGT TCCTGTAATC GCTGGTTCCTG CTTTGAAAGC      200
TCTTGAAGGC GATGCTTCAT ACGAAGAAAA AATCATGGAA TTAATGGCTG      250
CAGTTGACGA ATACGTTCCA ACTCCAGAAC GTGACACTGA CAAACCATTC      300
ATGATGCCAG TCGAAGACGT ATTCTCAATC ACTGGACGTG GTACTGTTGC      350
TACAGGCCGT GTTGAACGTG GACAAGTTCG CGTTGGTGAC GAAGTTGAAA      400
TCGTTGGTAT TGCTGAAGAA ACTGCTAAAA CAACTGTAAC TGGTGTGAA      450
ATGTTCCGTA AATTGTTAGA CTATGCTGAA GCAGGGGATA ACATTGGTGC      500
ATTGCTACGT GGTGTTGCTC GTGAAGACAT CCAACGTGGA CAAGTATTGG      550
CTAAAGCTGG TACAATCACA CCTCATACAA AATTAAAGC TGAAGTTTAC      600
GTTTTAACAA AAGAAGAAGG TGGACGTCAC ACACCATTCT TC          642

```

## 2) INFORMATION FOR SEQ ID NO: 882

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 636 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus casseliflavus*  
 (B) STRAIN: R775

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 882

20	CCTATGCCTC	AAACACGTGA	ACACATCTTG	TTATCACGTA	ACGTTGGTGT	50
	ACCATACATC	GTTGTTTTCT	TAAACAAAAT	GGATATGGTT	GATGACGAAG	100
	AATTACTAGA	ATTAGTTGAA	ATGGAAGTTC	GTGACTTATT	GTCAGAATAT	150
	GACTTCCCAG	GCGACGATGT	TCCTGTAATC	GCTGGTTCTG	CTTTGAAAGC	200
	TCTTGAAGGC	GATGCTTCAT	ACGAAGAAAA	AATCATGGAA	TTAATGGCTG	250
25	CAGTTGACGA	ATACGTTCCA	ACTCCAGAAC	GTGACACTGA	CAAACCATTC	300
	ATGATGCCAG	TCAAGACGT	ATTCTCAATC	ACTGGACGTG	GTAAGTTGAA	350
	TACAGGCCGT	GTTGAACGTG	GACAAGTTCG	CGTTGGTGAC	GAAGTTGAAA	400
	TCGTTGGTAT	TGCTGAAGAA	ACTGCTAAAA	CAACTGTAAC	TGGTGTGAA	450
	ATGTTCCGTA	AATTGTTAGA	CTATGCTGAA	GCAGGGGATA	ACATTGGTGC	500
30	ATTGCTACGT	GGTGTGCTC	GTGAAGACAT	CCAACGTGGA	CAAGTATTGG	550
	CTAAAGCTGG	TACAATCACA	CCTCATACAA	AATTTAAAGC	TGAAGTTTAC	600
	GTTTTAACAA	AAGAAGAAGG	TGGACGTCAT	ACACCA		636

## 2) INFORMATION FOR SEQ ID NO: 883

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 641 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*  
 (B) STRAIN: R422

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 883

55	GGTCCTATGC	CTCAAACACG	TGAACATATC	TTATTATCAC	GTAACGTTGG	50
	TGTACCATAC	ATCGTTGTAT	TCTTAAACAA	AATGGATATG	GTTGATGACG	100
	AAGAATTATT	AGAATTAGTA	GAAATGGAAG	TTCGTGACTT	ATTATCAGAA	150
	TACGATTTCC	CAGGCGATGA	TGTTCCAGTT	ATCGCAGGTT	CTGCTTTGAA	200
	AGCTTTAGAA	GGCGACGAGT	CTTATGAAGA	AAAAATCTTA	GAATTAATGG	250
	CTGCAGTTGA	CGAATATATC	CCAACCTCCAG	AACGTGATAC	TGACAAACCA	300
	TTCATGATGC	CAGTCGAAGA	CGTATTCTCA	ATCACTGGAC	GTGGTACTGT	350
	TGCTACAGGC	CGTGTGTAAC	TGCGTGAAGT	TCGCGTTGGT	GACGAAGTTG	400
60	AAATCGTTGG	TATTAAAGAC	GAAACATCTA	AAACAACGTG	TACAGGTGTT	450

GAAATGTTCC	GTAAATTATT	AGACTACGCT	GAAGCAGGCG	ACAACATCGG	500
TGCTTTATTA	CGTGGTGTAG	CACGTGAAGA	TATCGAACGT	GGACAAGTAT	550
TAGCTAAACC	AGCTACAATC	ACTCCACACA	CAAATTCAA	AGCTGAAGTA	600
TACGTATTAT	CAAAAGAAGA	AGGCGGACGT	CACACTCCAT	T	641

5

## 2) INFORMATION FOR SEQ ID NO: 884

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 640 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear

15

- (ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*
- (B) STRAIN: R575

20

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 884

TATGCCTCAA	ACACGTGAAC	ATATCTTATT	ATCACGTAAC	GTTGGTGTAC	50
CATACATCGT	TGTATTCTTA	AACAAAATGG	ATATGGTTGA	TGACGAAGAA	100
TTATTAGAAT	TAGTAGAAAT	GGAAGTTCGT	GACTTATTAT	CAGAATACGA	150
TTTCCCAGGC	GATGATGTTT	CAGTTATCGC	AGGTTCTGCT	TTGAAAGCTT	200
TAGAAGGCCA	CGAGTCTTAT	GAAGAAAAAA	TCTTAGAATT	AATGGCTGCA	250
GTTGACGAAT	ATATCCCAAC	TCCAGAACGT	GATACTGACA	AACCATTTCAT	300
30 GATGCCAGTC	GAAGACGTAT	TCTCAATCAC	TGGACGTGGT	ACTGTTGCTA	350
CAGGCCGTGT	TGAACGTGGT	GAAGTTCGCG	TTGGTGACGA	AGTTGAAATC	400
GTTGGTATTA	AAGACGAAAC	ATCTAAAACA	ACTGTTACAG	GTGTTGAAAT	450
GTTCCGTAAA	TTATTAGACT	ACGCTGAAGC	AGGCGACAAC	ATCGGTGCTT	500
TATTACGTGG	TGTTGCACGT	GAAGATATCG	AACGTGGACA	AGTATTAGCT	550
35 AAACCAGCTA	CAATCACTCC	ACACACAAAA	TTCAAAGCTG	AAGTATACGT	600
ATTATCAAAA	GAAGAAGGCG	GACGTCACAC	TCCATTCTTC		640

## 40 2) INFORMATION FOR SEQ ID NO: 885

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 632 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear

45

- (ii) MOLECULE TYPE: Genomic DNA

- (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*
- (B) STRAIN: R492

50

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 885

55 TGCCTCAAAC	TCGTGAACAC	ATCCTATTGT	CTCGTCAAGT	TGGTGTTTCCT	50
TACATCGTTG	TATTCTTGAA	CAAAGTAGAC	ATGGTTGATG	ACGAAGAATT	100
ACTAGAATTA	GTTGAAATGG	AAGTTCGTGA	CCTATTAACA	GAATACGAAT	150
TCCCTGGTGA	CGATGTTTCT	GTAAGTTGCTG	GATCAGCTTT	GAAAGCTCTA	200
60 GAAGGCGACG	CTTCATACGA	AGAAAAAATT	CTTGAATTAA	TGGCTGCAGT	250

700

	TGACGAATAC	ATCCCAACTC	CAGAACGTGA	CAACGACAAA	CCATTCATGA	300
	TGCCAGTTGA	AGACGTGTTT	TCAATTACTG	GACGTGGTAC	TGTTGCTACA	350
	GGTCGTGTTG	AACGTGGACA	AGTTCGCGTT	GGTGACGAAG	TTGAAGTTGT	400
	TGGTATTGCT	GAAGAACTT	CAAAAACAAC	AGTTACTGGT	GTTGAAATGT	450
5	TCCGTAAATT	GTTAGACTAC	GCTGAAGCTG	GAGACAACAT	TGGTGCTTTA	500
	CTACGTGGTG	TTGCACGTGA	AGACATCCAA	CGTGGACAAG	TTTGTAGCTAA	550
	ACCAGGTACA	ATCACACCTC	ATACAAAATT	CTCTGCAGAA	GTATACGTGT	600
	TGACAAAAGA	AGAAGGTGGA	CGTCATACTC	CA		632

10

## 2) INFORMATION FOR SEQ ID NO: 886

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 640 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: R576

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 886

	CGGCCCAATG	CCTCAAACCTC	GTGAACACAT	CCTATTGTCT	CGTCAAGTTG	50
	GTGTTCTCTTA	CATCGTTGTA	TTCTTGAACA	AAGTAGACAT	GGTTGATGAC	100
30	GAAGAATTAC	TAGAATTAGT	TGAAATGGAA	GTTCGTGACC	TATTAACAGA	150
	ATACGAATTC	CCTGGTGACG	ATGTTCTCTGT	AGTTGCTGGA	TCAGCTTTGA	200
	AAGCTCTAGA	AGGCGACGCT	TCATACGAAG	AAAAAATTCT	TGAATTAATG	250
	GCTGCAGTTG	ACGAATACAT	CCCAACTCCA	GAACGTGACA	ACGACAAACC	300
	ATTCATGATG	CCAGTTGAAG	ACGTGTTCTC	AATTACTGGA	CGTGGTACTG	350
35	TTGCTACAGG	TCGTGTTGAA	CGTGGACAAG	TTCGCGTTGG	TGACGAAGTT	400
	GAAGTTGTTG	GTATTGCTGA	AGAACTTCA	AAAACAACAG	TTACTGGTGT	450
	TGAAATGTTT	CGTAAATTGT	TAGACTACGC	TGAAGCTGGA	GACAACATTG	500
	GTGCTTTACT	ACGTGGTGTG	GCACGTGAAG	ACATCCAACG	TGGACAAGTT	550
	TTAGCTAAAC	CAGGTACAAT	CACACCTCAT	ACAAAATTCT	CTGCAGAAGT	600
40	ATACGTGTTG	ACAAAAGAAG	AAGGTGGACG	TCATACTCCA		640

## 2) INFORMATION FOR SEQ ID NO: 887

45

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 806 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 50 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 55 (A) ORGANISM: *Zoogloea ramigera*  
 (B) STRAIN: ATCC 25935

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 887

60	ATCCTGGTTT	GCTCCGCAGC	TGACGGCCCA	ATGCCACAGA	CCCGCGAGCA	50
----	------------	------------	------------	------------	------------	----

	CATCCTGCTG	CCCCGCCAAG	TTGGCGTTCC	ATACATCATC	GTGTTCTCTGA	100
	ACAAGTGCGA	CCTGGTTGAC	GACGCAGAAC	TGCTGGAAC	GGTCGAAATG	150
	GAAGTGCGTG	AATTGCTGTC	GAAATACGAG	TTCCCAGGCG	ACGACGTACC	200
	AATCATCAAG	GGTTCGGCAC	GTATGGCGCT	GGAAGGCAAA	GAAGGCGAGA	250
5	TGGGCGTTGA	CGCCATCATG	CGTCTGGCCG	ATGCACTGGA	CAGCTACATC	300
	CCTACGCCAG	AGCGCGCAGT	CGATGGCGCC	TTCTTGATGC	CAGTGGAAGA	350
	CGTGTTCTCG	ATCTCGGGTC	GCGGTACCGT	TGTGACCGGT	CGTATCGAGC	400
	GCGGCGTGAT	CAAGGTCGGC	GAAGAGATCG	AAATCGTCGG	CATTATCGAC	450
	ACCGTCAAAA	CCACTTGCAC	CGGCGTGGA	ATGTTCCGCA	AGCTGCTGGA	500
10	CCAGGGTCAA	GCCGGCGACA	ACGTTGGTCT	GCTGCTGCGC	GGCACCAAGC	550
	GTGAAGACGT	ACAGCGTGGT	CAGGTTCTGG	CCAAGCCAGC	GTCGATCAAG	600
	CCGCACAACC	ACTTCACCGG	CGAGATCTAC	GTTCTGTCTGA	AAGATGAAGG	650
	CGGCCGTCAC	ACCCCGTTCT	TCAACAAC	TCGTCCACAG	TTCTACTTCC	700
	GTACGACTGA	CGTGACCGGT	TCGATCGAAC	TGCCAGCAGA	CAAAGAAATG	750
15	GTCATGCCAG	GCGACAACGT	GTCGATCACC	GTCAAGCTGA	TCAACCCGAT	800
	CGCGAT					806

## 20 2) INFORMATION FOR SEQ ID NO: 888

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 634 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*  
 (B) STRAIN: R503

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 888

35	TATGCCTCAA	ACACGTGAAC	ATATCTTATT	ATCACGTAAC	GTTGGTGTAC	50
	CATACATCGT	TGTATTCTTA	AACAAAATGG	ATATGGTTGA	TGACGAAGAA	100
	TTATTAGAA	TAGTAGAAAT	GGAAGTTTCG	GACTTATTAT	CAGAAATACGA	150
	TTTCCCAGGC	GATGATGTTT	CAGTTATCGC	AGGTTCTGCT	TTGAAAGCTT	200
40	TAGAAGGCGA	CGAGTCTTAT	GAAGAAAAAA	TCTTAGAATT	AATGGCTGCA	250
	GTTGACGAAT	ATATCCCAAC	TCCAGAACGT	GATACTGACA	AACCATTTCAT	300
	GATGCCAGTC	GAAGACGTAT	TCTCAATCAC	TGGACGTGGT	ACTGTTGCTA	350
	CAGGCCGTGT	TGAACGTGGT	GAAGTTCGCG	TTGGTGACGA	AGTTGAAATC	400
	GTTGGTATTA	AAGACGAAAC	ATCTAAAAAC	ACTGTTACAG	GTGTTGAAAT	450
45	GTTCCGTAAA	TTATTAGACT	ACGCTGAAGC	AGGCGACAAC	ATCGGTGCTT	500
	TATTACGTGG	TGTAGCACGT	GAAGATATCG	AACGTGGACA	AGTATTAGCT	550
	AAACCAGCTA	CAATCACTCC	ACACACAAAA	TTCAAAGCTG	AAGTATACGT	600
	ATTATCAAAA	GAAGAAGGCG	GACGTCACAC	TCCA		634

50

## 2) INFORMATION FOR SEQ ID NO: 889

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 493 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus fumigatus*  
 (B) STRAIN: ATCC 14110

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 889

	TGTCTTCATC	CAGGAATTGA	TTGTGAGTCG	TTCCACATGC	TCACCTAGTT	50
	TTCGCTCGAT	CTTTTCACTA	ACGCAAACCA	TGTAGAACAA	CATTGCCAAG	100
10	GCCCACGGTG	GTTACTCCGT	CTTCACTGGT	GTTGGTGAGC	GTA CTCTGTA	150
	GGGTAACGAT	CTGTACCACG	AAATGCAGGA	GACTGGTGTC	ATTCAGCTCG	200
	AGGGTGAATC	CAAGGTCGCA	CTGGTGTTTCG	GACAGATGAA	CGAGCCCCCC	250
	GGTGCCCGTG	CCCGTGTCGC	CCTTACCGGT	CTGACCATTG	CCGAGTACTT	300
	CCGTGACGAG	GAGGGTCAGG	ACGTGCTGCT	CTTCATTGAC	AACATTTTCC	350
15	GTTTCACCCA	GGCCGGTTCT	GAGGTGTCTG	CCCTTCTCGG	TCGTATCCCC	400
	TCTGCCGTCG	GTTACCAGCC	CACCCTGGCC	GTCGACATGG	GTGGTATGCA	450
	GGAGCGTATC	ACCACCACCA	AGAAGGGTTC	TATTACCTCC	GTC	493

## (i) SEQUENCE CHARACTERISTICS: 890

- (A) LENGTH: 466 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Penicillium marneffei*  
 (B) STRAIN: ATCC 64101

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 890

35	TTGATTGTAC	GTCTTTACCT	TTCTGCCTGA	CTGTTTACGA	CAACTAACGA	50
	AAGCGTAGAA	CAACATTGCC	AAGGCTCACG	GTGGTTACTC	TGTCTTCACT	100
	GGTGTGCGTG	AACGTACTCG	TGAGGGTAAC	GATTTGTACC	ACGAAATGCA	150
	GGAAACTGGT	GTCATTACAGC	TCGAGGGTGA	ATCCAAGGTC	GCCCTCGTGT	200
	TCGGTCAGAT	GAACGAGCCC	CCCGGTGCC	GTGCCCGTGT	CGCTCTTACT	250
40	GGTTTGACCA	TTGCCGAGTA	CTTCCGTGAC	GAGGAAGGTC	AGGACGTGCT	300
	TCTCTTCATT	GACAACATT	TCCGTTTCAC	TCAGGCCGGT	TCTGAGGTGT	350
	CTGCCCTTCT	GGGTCTGATC	CCCTCTGCCG	TCGGTTACCA	GCCCACCCTT	400
	GCCGTGACCA	TGGGTATCAT	GCAGGAGCGT	ATTACCACCA	CCACCAAGGG	450
45	TTCCATCACC	TCCGTC				466

## 2) INFORMATION FOR SEQ ID NO: 891

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 478 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Paecilomyces lilacinus*  
 (B) STRAIN: ATCC 42570

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 891

```

5  AGGAGCTGAT CAACAACATC GCCAAGGCTC ACGGTGGTTA CTCCGTCTTC      50
   ACTGGTGTCTG GTGAGCGTAC CCGTGAGGGT AACGATCTGT ACCACGAAAT      100
   GCAGGAGACC TCGGTCATTC AGCTCGAGGG CGAGTCTAAG GTGGCCCTGG      150
   TCTTTGGTCA GATGAACGAG CCCCCGGGTG CTCGTGCCCC TGTCGCTCTT      200
   ACTGGTCTTA CCGTCGCCGA GTACTTCCGT GACCAGGAGG GTCAGGATGG      250
   TTAGTTCTCG TCCACTCATG CCGAAACATG TGC GTGTTCC GAGGCTAATC      300
10 AACGTGCCAG TGCTGCTTTT CATCGACAAC ATTTTCCGAT TCACACAGGC      350
   CGGTTCCGAG GTGTCTGCCC TGCTGGGTCG TATCCCCCTCT GCCGTCGGTT      400
   ACCAGCCAC CCTCGCCGTC GACATGGGTG GCATGCAGGA GCGTATCACC      450
   ACCACCAAGA AGGGCTCTAT CACCTCCG      478

```

15

## 2) INFORMATION FOR SEQ ID NO: 892

```

20 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 481 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

```

25 (ii) MOLECULE TYPE: Genomic DNA

```

```

    (vi) ORIGINAL SOURCE:

```

```

    (A) ORGANISM: Penicillium marneffe
    (B) STRAIN: ATCC 58950

```

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 892

```

35 GTCTTTATCC AGGAGTTGAT TGTACGTCTT TACCTTTCTG CCTGACTGTT      50
   TACGACAAC TACGAAAGCG TAGAACAACA TTGCCAAGGC TCACGGTGGT      100
   TACTCTGTCT TCACTGGTGT CCGTGAACGT ACTCGTGAGG GTAACGATTT      150
   GTACCAGGAA ATGCAGGAAA CTGGTGTCTAT TCAGCTCGAG GGTGAATCCA      200
   AGGTCCGCCCT CGTGTTCGGT CAGATGAACG AGCCCCCGG TGCCCGTGCC      250
   CGTGTCTGCTC TTA CTGGTTT GACCATTGCC GAGTACTTCC GTGACGAGGA      300
   AGGTCAGGAC GTGCTTCTCT TCATTGACAA CATTTTCCGT TTCACTCAGG      350
40 CCGGTTCTGA GGTGTCTGCC CTTCTGGGTC GTATCCCCTC TGCCGTCGGT      400
   TACCAGCCCA CCCTTGCCGT CGACATGGGT ATCATGCAGG AGCGTATTAC      450
   CACCACCACC AAGGGTTCCA TCACCTCCGT C      481

```

45

## 2) INFORMATION FOR SEQ ID NO: 893

```

50 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1208 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

```

55 (ii) MOLECULE TYPE: Genomic DNA

```

```

    (vi) ORIGINAL SOURCE:

```

```

    (A) ORGANISM: Sporothrix schenckii
    (B) STRAIN: ATCC 14285

```

```

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 893

```

	CTCGTCCGTG	GTGCTAAGGC	CACTGACACT	GGTTSCCCCA	TTACCATCCS	50
	CGTCGGCCCC	GGTACCCTCG	GTGCGATCAT	GAACGTCACC	GGTGACCCGA	100
	TCGACGAGCG	CGGTCCCATC	AAGACCGACA	AGTTCGGTCC	CATCCACGCT	150
5	GAGGCTCCCG	AGTTCGTTGA	CCAGTCGACC	ACCGCTGAGG	TTCTCGTGAC	200
	TGGTATCAAG	GTCGTCGATC	TGCTTGCTCC	CTACGCCCGT	GGTGGTAAGA	250
	TTGGTCTGTT	TGGCGGTGCC	GGTGTGGCA	AGACCGTGTT	CATCCAGGAG	300
	CTCATCAACA	ACATCGCCAA	GGCCCACGGT	GGTTACTCCG	TCTTCACCGG	350
	TGTCGGCGAG	CGTACCCGTG	AGGGTAACGA	TCTGTACCAC	GAAATGCAGG	400
10	AGACCTCTGT	CATTTCAGCTT	GACGGTGACT	CCAAGGTCGC	CCTGGTGTTT	450
	GGTCAGATGA	ACGAGCCCCC	TGGTGCTCGT	GCCCGTGTCG	CCTTGACCGG	500
	TTTGACTGTC	GCTGAGTACT	TCCGTGACGA	GGAGGGCCAG	GATGGTATGT	550
	TTTGAATTAT	TTCTTTGTCG	TACAGTTCCA	AATCGAAGAA	TACTAACTT	600
	GTCAGTGCTT	CTCTTCATCG	ACAACATTTT	CCGCTTCACC	CAGGCCGGTT	650
15	CTGAGGTGTC	TGCCCTTCTG	GGTCGTATT	CCTCCGCTGT	CGGTTACCAG	700
	CCCACGCTCG	CCGTGGACAT	GGGTCTGATG	CAGGAGCGTA	TTACCACCAC	750
	CCGCAAGGGC	TCAATTACCT	CCGTCCAGGC	CGTCTACGTG	CCCGCTGACG	800
	ATCTGACGGA	TCCCGCCCCC	GCCACCACCT	TCGCCCATCT	GGACGCCACC	850
	ACTGTGCTGT	CCCGAGGTAT	CTCTGAGCTG	GGTATCTACC	CCGCTGTCTGA	900
20	CCCCCTCGAC	TCCAAGTCGC	GTATGCTGGA	CCCCCGTATT	GTCGGTGACG	950
	ACCACTACGA	GACCGCCACT	CGCGTCCAGT	AGATCCTCCA	AGAGTACAAG	1000
	TCGCTGCAGG	ACATCATCGC	CATTCTGGGT	ATGGACGAGC	TGTCTGAGGC	1050
	CGACAAGCTT	ACAGTCGAGC	GTGCTCGTAA	GATCCAGCGT	TTCCTGAGCC	1100
	AGCCGTTTAC	GGTCGCGCAG	GTCTTCACTG	GTATCGAAGG	CCAGCTGGTC	1150
25	GATCTGAAGG	ACACTATCGC	TTCGTTCAAG	GCTATCCTGA	GCGGTGAGGG	1200
	TGACAGCC					1208

## 30 2) INFORMATION FOR SEQ ID NO: 894

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 534 bases  
 (B) TYPE: Nucleic acid  
 35 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 40 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Malbranchea filamentosa*  
 (B) STRAIN: ATCC 48174

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 894

45	TTCAGGAAGT	TATTGTAAGC	CGCCCTCTTT	ATGCATTGAG	GGTGAATAAG	50
	AAGGCTGACA	GGTAATAGAA	CAACATTGCC	AAGGCTCACG	GTGGTTACTC	100
	CGTGTTCACT	GGTGTCGGTG	AGCGTACCCG	TGAAGGAAAC	GATCTGTACC	150
	ATGAAATGCA	GGAAACCCGC	GTCATCCAGC	TTGATGGCGA	GTCTAAGGTC	200
50	GCCCTTGTGT	TCGGTCAGAT	GAACGAGCCC	CCTGGAGCCC	GTGCCCCTGT	250
	CGCTCTTACT	GGTCTTACCG	TTGCCGAATA	CTTCCGTGAC	GAGGAGGGCC	300
	AAGATGGTAC	GCCTTTTTTAC	TCTTCTTATT	CTTCGGGTGCG	GACTACAGAA	350
	CTAACCTGCT	CCAGTGCTTC	TCTTCATTGA	TAACATTTTC	CGTTTCACAC	400
	AAGCCGGTTC	TGAGGTGTCT	GCCTTGCTTG	GACGTATTCC	CTCTGCCGTC	450
55	GGTTACCAGC	CCACTCTCGC	CGTCGACATG	GGTGGTATGC	AGGAACGTAT	500
	CACAACCACC	AACAAGGGTT	CCATTACTTC	CGTG		534

## 60 2) INFORMATION FOR SEQ ID NO: 895



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 448 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Paecilomyces lilacinus*  
 (B) STRAIN: ATCC 60735

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 895

CAAGGCTCAC	GGTGGTTACT	CCGTCTTCAC	TGGTGTCGGT	GAGCGTACCC	50
GTGAGGGTAA	CGATCTGTAC	CACGAAATGC	AGGAGACCTC	GGTCATTACG	100
CTCGAGGGCG	AGTCTAAGGT	GGCCCTGGTC	TTTGGTCAGA	TGAACGAGCC	150
CCCGGGTGCT	CGTGCCCGTG	TCGCTCTTAC	TGGTCTTACC	GTCGCCGAGT	200
ACTTCCGTGA	CCAGGAGGGT	CAGGATGGTT	AGTTCTCGTC	CACTCATGCC	250
GAAACATGTG	CGTGTTCCGA	GGCTAATCAA	CGTGCCAGTG	CTGCTTTTCA	300
TCGACAACAT	TTTCCGATTC	ACCCAGGCCG	GTTCCGAGGT	GTCTGCCCTG	350
CTGGGTCGTA	TCCCCTCTGC	CGTCGGTTAC	CAGCCCACCC	TCGCCGTCGA	400
CATGGGTGGC	ATGCAGGAGC	GTATCACCAC	CACCAAGAAG	GGCTCTAT	448

## 2) INFORMATION FOR SEQ ID NO: 896

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 483 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus niger*  
 (B) STRAIN: ATCC 9508

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 896

ATTCAGGAGC	TGATTGTAAG	TTGCCAATCC	ATGAACTGGA	GATTTGGTGT	50
GACCCATAGA	ACTAACAAAT	TATTTAGAAC	AACATCGCCA	AGGCTCACGG	100
TGGTTACTCC	GTCTTCTGTG	GTGTCGGTGA	GCGTACTCGT	GAGGGTAACG	150
ATCTGTACCA	CGAAATGCAG	GAGACTGGTG	TCATCCAGCT	CGAGGGTGAC	200
TCCAAGGTCG	CTCTGGTCTT	CGGTCAGATG	AACGAGCCCC	CGGGTGCCCC	250
TGCCCCGTGTC	GCCCTTACCG	GTCTGACCAT	TGCCGAGTAC	TTCCGTGACG	300
AGGAGGGTCA	GGACGTGCTG	CTCTTCATTG	ACAACATTTT	CCGTTTCACC	350
CAGGCCGGTT	CTGAGGTGTC	TGCCCTTCTG	GGTCGTATCC	CCTCTGCCGT	400
CGGTTACCAG	CCCACTCTGG	CCGTCGACAT	GGGTGGTATG	CAGGAGCGTA	450
TTACCACCAC	CACCAAGGGT	TCCATTACCT	CCG		483

## 2) INFORMATION FOR SEQ ID NO: 897

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1124 bases

(B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus fumigatus*  
 (B) STRAIN: ATCC 14110

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 897

	AAGCTCAAGT	CCGAGCGTGA	GCGTGGTATC	ACCATCGACA	TTGCCCTCTG	50
	GAAGTTCCAG	ACTCCCAAGT	ATGAGGTCAC	TGTCATCGGT	AAGCTCGACT	100
15	CGCCCCGATA	TGTTTTGGTG	CTGTAGCTAA	CACGATCTGA	AGATGCCCCC	150
	GGTCACCGTG	ACTTCATCAA	GAACATGATC	ACTGGTACCT	CCCAGGCTGA	200
	CTGCGCTATC	CTCATCATTTG	CCTCCGGTAC	TGGTGAGTTC	GAGGCTGGTA	250
	TCTCCAAGGA	TGGCCAGACC	CGTGAGCACG	CTCTGCTGGC	TTTCACCCTC	300
	GGTGTCAAGC	AGCTCATCGT	CGCCCTCAAC	AAGATGGACA	CCTGCAAGTG	350
20	GTCCGAGGAT	CGTTACAACG	AAATTGTCAA	GGAAACCTCC	AACTTCATCA	400
	AGAAGGTCGG	CTACAACCCC	AAGGCCGTTT	CCTTCGTCCC	CATCTCTGGC	450
	TTCAACGGTG	ACAACATGCT	TGAGCCCTCC	TCCAACCTGCC	CCTGGTACAA	500
	GGGATGGGAG	AAGGAGACCA	AGGCCGGCAA	GGTCACTGGT	AAGACCCTCA	550
	TCGAGGCCAT	CGACGCCATT	GAGCCCCCTG	TCCGTCCCTY	CAACAAGCCC	600
25	CTCCGTCTTC	CCCTCCAGGA	TGTCTACAAG	ATCTCTGGTA	TCGGAACGGT	650
	CCCTGTCGGC	CGTGTGAGAG	CCGGTATCAT	CAAGCCCGGC	ATGGTTCGTCA	700
	CCTTCGCCCC	CGCCAACGTC	ACCACTGAAG	TCAAGTCCGT	CGAGATGCAC	750
	CACCAGCAGC	TCCAGGAGGG	TGTCCCCGGT	GACAACGTCG	GTTTCAACGT	800
	CAAGAACGTT	TCCGTCAAGG	AAGTCCGCCG	TGGTAACGTC	TGCGGTGACT	850
30	CCAAGAACAG	TCCCCCTCAG	GGTGCTGCCT	CCTTCAACGC	CCAGGTCATC	900
	GTCCTCAACC	ACCCCGGTCA	GGTCGGCGCT	GGTTACGCCC	CCGTCCTCGA	950
	CTGCCACACT	GCCCACATTG	CTTGCAAGTT	CTCTGAGCTG	CTTGAGAAGA	1000
	TTGACCGCCG	TACCGGCAAG	TCTGTTGAGA	ACAACCCCAA	GTTTCATCAAG	1050
	TCCGGTGATG	CCGCCATCGT	GAAGATGGTT	CCTTCCAAGC	CCATGTGTGT	1100
35	CGAGTCCTTC	ACTGACTACC	CCCC			1124

2) INFORMATION FOR SEQ ID NO: 898

40

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1363 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Penicillium marneffe*  
 (D) STRAIN: WSA-214

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 898

55	CTCAAGGCTG	AGCGTGAGCG	TGGTATCACC	ATCGATATTG	CTCTCTGGAA	50
	GTTCCAGACT	GCCAAGTACG	AGGTTACCGT	CATTGACGCC	CCCGGTCACC	100
	GTGATTTTCA	CAAGAACATG	ATCACTGGTA	CCTCCCAGGC	CGATTGCGCT	150
	ATTCTCATCA	TTGCCCTCTG	TACTGGTGAA	TTCGAGGCTG	GTATCTCCAA	200
	GGATGGCCAG	ACTCGTGAGC	ACGCTCTTTT	GGCTTTTACC	CTCGGTGTCC	250
60	GTCAGCTCAT	TGTTGCCCTC	AACAAGATGG	ACACTTGCAA	GTGGTCTCAG	300

```

GGTGAGTACT CGTACCTGCG TTTGGCCTTG AATATCTTAC TAATGCACCA 350
TAGATCGTTA CAACGAAATT GTCAAGGAGA CTTCCAACCTT CATCAAGAAG 400
GTCGGATACA ACCCTAAGAA CGTTCCCTTTT GTTCCTATCT CCGGTTTCAA 450
CGGTGACAAC ATGCTTGAGC CCTCCCCCAA CTGCCCTTGG TACAAGGGTT 500
5 GGGAGAAGGA GACCAAGGCC GGTAAGGTCA CTGGTAAGAC CCTCCTCGAG 550
GCCATCGACG CCATTGAGCC CCCTACCCGT CCCGCCAACA AGGTTAGTCC 600
CTCCTCGACT ACTCAAACCC TCCTCATAAG TTCATGATTA CGACTCGTTC 650
ACAGCCCCTC CGTCTTCCCC TCCAGGATGT CTACAAGATC GGTGGTATTG 700
GAACGGTTCC CGTCGGTCGT GTTGAGACTG GTACCATCGT TCCTGGTATG 750
10 GTTGTACACT TGTAAGTCAC TCTCCTCGCT TATCCTACCT GAAATCATCA 800
TGTGCTAACT TGACACTCAG CGCTCCCGCC AACGTCACCA CTGAAGTCAA 850
GAGTGTTGAA ATGCACCACC AGCAGCTCAC TGCCGGTCAG CCCGGTGACA 900
ACGTTGGTTT CAACGTGAAG AACGTCTCCG TCAAGGAAAT CCGTCGTGGT 950
AACGTTGCTG GTGACAGCAA GAACGACCCC CCTGCCGGTG CTGCCCTCCTT 1000
15 CAACGCCCAG GTCATCGTCC TCAACCACCC CGGTCAGGTC GGTGCTGGTT 1050
ACGCCCCAGT CCTCGATTGC CACACTGCCC ACATTGCTTG CAAGTTCGCT 1100
GAGCTCCTCG AGAAGATTGA CCGTCGTACC GGAAAGTCTG TTGAGGACCA 1150
CCCCAAGTTC ATCAAGTCCG GTGACGCTGC CATCGTCAAG ATGATTCCTT 1200
CCAAGCCCAT GTGTGTTGAG GCTTTCACCG AGTACCCTCC TCTCGGTCGT 1250
20 TTCGCCGTTT GCGAGTAAGT TTTATCTCCG TTGTCTATTT TCCATCCTTC 1300
CCTTCTCCTC CGTCTTCCAT ATATATTTTT TCAGTTATAT GTGACTAACC 1350
ACAAATCACG GGA 1363

```

25

## 2) INFORMATION FOR SEQ ID NO: 899

## (i) SEQUENCE CHARACTERISTICS:

```

30 (A) LENGTH: 1147 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

35

## (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Piedraia hortai
    (B) STRAIN: ATCC 24292

```

## 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 899

```

CAAGCTGAAA GCCGAGCGTG AGCGTGGTAT CACTATCGAC ATTGCCCTCT 50
GGAAGTTCGA GACTCCCAAG TACTATGTCA CCGTCATTGG TACGTCGCAT 100
TATCTCACTC CTCACAGAAG CACGCTCCTA ACATCACACA GACGCTCCCG 150
45 GTCACCGTGA TTTCATCAAG AACATGATCA CTGGTACCTC CCAGGCCGAC 200
TGCGCTATCC TCATTATCGC TGCCGGTACT GGTGAGTTCG AGGCTGGTAT 250
CTCCAAGGAT GGCCAGACTC GTGAGCACGC CCTGCTCGCC TACACCCTCG 300
GTGTGCGTCA GCTCATCGTC GCCATCAACA AGATGGACAC CACCAAGTGG 350
TCTGAGGCCG GTTACCAGGA GATCATCAAG GAGACCTCCA ACTTCATCAA 400
50 GAAGGTCGGC TACAACCCCA AGACCGTCGC TTTCGTCCCC ATCTCTGGCT 450
TCAACGGCGA CAACATGCTT GCCCCCTCCA CCAACTGCCC CTGGTACAAG 500
GGATGGGAGC GTGAGGTCAA GGGCAACAAG CAGACCGGCA AGACCCTCCT 550
CGAGGCCATT GACGGCATTG AGCCCCCAA GCGTCCCTCC GACAAGCCCC 600
TCCGTCTTCC TCTCCAGGAT GTTTACAAGA TCGGTGGTAT CGGAAGCTGT 650
55 CCTGTCGGCC GTATCGAGAC TGGTGTCTTC AAGCCCGGTA TGGTCTGTAC 700
CTTCGCTCCC GCCAACGTCA CCACTGAAGT CAAGTCCGTC GAGATGCACC 750
ACGAGCAGCT CACTGAGGGT CTTCCCGGTG ACAACGTCGG TTTCAACGTG 800
AAGAAGCTTT CCGTCAAGGA CATCCGCGGT GGTAACGTTG CCAGTGACTC 850
CAAGAACGAC CCCGCTCTGG GTGCCGCTTC TTTCGACGCC CAGGTCATCG 900
60 TCCTCAACCA CCCCGGTCAG GTCGGTGCTG GTTACGCCCC GGTCCCTCGAT 950

```

TGCCACACTG	CCCACATTGC	TGCAAGTTC	TCCGAGATCA	AGGAGAAGAT	1000
CGACCGCCGT	ACCGGCAAGT	CTGTTGAGGA	CGCCCCCAAG	TTCATCAAGT	1050
CTGGTGACTC	TGCCATCGTC	AAGATGGTTC	CCTCCAAGCC	CATGTGCGTT	1100
GAGGCTTTCA	CCGACTACCC	TCCTCTGGGC	CGTTTCGCCG	TCCGTGA	1147

5

## 2) INFORMATION FOR SEQ ID NO: 900

- 10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1150 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 15 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Paecilomyces lilacinus*  
 (B) STRAIN: ATCC 60735
- 20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 900

CTCAAGGCCG	AGCGTGAGCG	TGGTATCACC	ATCGACATTG	CCCTCTGGAA	50
GTTCGAGACT	CCCAAGTACT	ATGTCACCGT	CATTGGTACG	TCGACTCGCG	100
CGAGACTGGT	CGCAATTTCC	ACGTCGCTAA	CGTGCTTGAA	CAGACGCTCC	150
CGGCCACCGT	GACTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCTG	200
ACTGCGCTAT	CCTCATTATC	GCTGCCGGCA	CTGGTGAGTT	CGAGGCTGGT	250
ATCTCCAAGG	ATGGCCAGAC	CCGTGAGCAC	GCTCTGCTCG	CCTACACCCT	300
CGGTGTTAAG	CAGCTCATCG	TCGCTATCAA	CAAGATGGAC	ACCACCAAGT	350
GGTCTGAGGC	CCGTTTCCAG	GAGATCATCA	AGGAGACCTC	CAACTTCATC	400
AAGAAGGTCG	GCTACAACCC	CAAGACCGTC	GCTTTCGTCC	CCATCTCTGG	450
TTTCCACGGC	GACAACATGC	TTTCCCCCTC	CACCAACTGC	CCCTGGTACA	500
AGGGCTGGGA	GAAGGAGACC	AAGGCTGGCA	AGTCCACCGG	CAAGACCCTC	550
CTTGAGGCCA	TCGACTCCAT	CGAGCCCCCC	AAGCGCCCCA	GCGACAAGCC	600
CCTCCGCTT	CCCCTTTCAGG	ATGTGTACAA	GATCGGCGGT	ATCGGCACAG	650
TCCCTGTTCG	CCGTATCGAG	ACTGGTGTCA	TCAAGCCCGG	CATGGTCGTG	700
ACCTTCGCTC	CTTCCAACGT	CACCACCGAA	GTCAAGTCCG	TTGAGATGCA	750
CCACGAGCAG	CTCTCCGAGG	GTGTCCCCGG	TGACAACGTC	GGCTTCAACG	800
TCAAGAACGT	CTCCGTCAAG	GAGATCCGTC	GTGGCAACGT	CGCCGGTGAC	850
TCCAAGAACG	ACCCCCCTCT	GGGTGCCGCT	TCTTTCGATG	CCCAGGTCAT	900
CGTCCTCAAC	CACCCCGGCC	AGGTCGGTGC	TGGCTACGCC	CCCGTCCTCG	950
ACTGCCACAC	CGCCCACATT	GCCTGCAAGT	TCGCCGAGAT	CAAGGAGAAG	1000
ATCGACCGCC	GTACCGGCAA	GTCTGTGCGAG	TCCGCCCCCA	AGTTCATCAA	1050
GTCTGGCGAC	TCTGCCATCG	TCAAGATGAT	TCCCTCCAAG	CCCATGTGCG	1100
TTGAGGCTTT	CACCGACTAC	CCTCCTCTGG	GCCGCTTCGC	CGTCCGTGAC	1150

## 50 2) INFORMATION FOR SEQ ID NO: 901

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 751 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 55
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- 60

(A) ORGANISM: *Paracoccidioides brasiliensis*  
 (B) STRAIN: ATCC 32075

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 901

5	TACCACTAAG	TGGTCCGAGA	CCCGATTCAA	TGAAATTATC	AAGGAAGTCA	50
	CCAATTTTCAT	TAAGAAGGTC	GGATATAACC	CCAAGACTGT	TCCTTTCGTT	100
	CCCATTTCTG	GTTTCGAGGG	TGACAACATG	ATCGAGCCCT	CTGCCAACTG	150
	CCCATGGTAC	AAGGGCTGGT	CCAAGGAGAC	TGCTCAGGGC	AAGTACTCTG	200
10	GCAAGACCCT	TCTTGAGGCC	ATCGACGCCA	TTGAGCCCCC	CACCCGTCCT	250
	ACCGATAAAC	CTCTCCGTCT	TCCCCTCCAG	GATGTCTACA	AGATCTCCGG	300
	TATTGGCACT	GTTCTGTGCG	GACGTGTTGA	GACTGGAGTC	ATCAAGCCCCG	350
	GTATGGTCGT	GACCTTCGCT	CCCGCCAACG	TCACCACTGA	AGTCAAGTCC	400
	GTCGAAATGC	ACCACCAGCA	GCTTACCGCC	GGTAACCCCG	GTGACAACGT	450
15	CGGCTTCAAC	GTCAAGAATG	TTTCCGTCAA	AGAAGTCCGC	CGTGGTAACG	500
	TTGCCGGTGA	CTCTAAGAA	GATCCCCCAA	AGGGCTGCGA	TTCCTTCAAT	550
	GCCCAGGTCA	TCGTCCTCAA	CCACCCCTGGT	CAGGTTGGCG	CTGGTTATGC	600
	CCCAGTCCTC	GACTGCCATA	CTGCCACAT	TGCCTGCAA	TTCGCTGAGC	650
	TCCTTGAGAA	GATTGATCGA	CGAACCGGAA	AGTCTGTTGA	GAACAACCCC	700
20	AAGTTCATCA	AGTCCGGTGA	TGCTGCTATC	GTCAAGATGA	TTCCTTCCAA	750
	G					751

25 2) INFORMATION FOR SEQ ID NO: 902

(i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 1056 bases
	(B) TYPE: Nucleic acid
30	(C) STRANDEDNESS: Double
	(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

35 (vi) ORIGINAL SOURCE:

	(A) ORGANISM: <i>Sporothrix schenckii</i>
	(B) STRAIN: ATCC 7968

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 902

40	CGTGAGCGCG	GTATCACCAT	CGATATTGCT	CTGTGGAAGT	TCGAGACCCC	50
	CAAGTACTAC	GTCACCGTCA	TTGACGCCCC	CGGTCATCGC	GATTTTCATCA	100
	AGAACATGAT	CACTGGTACC	TCGCAGGCCG	ACTGCGCCAT	TCTCATCATT	150
	GCCGCTGGTA	CTGGTGAGTT	CGAGGCTGGT	ATCTCCAAGG	ATGGCCAGAC	200
45	TCGTGAGCAC	GCTCTGCTCG	CCTACACCCT	GGGTGTGCGG	CAGCTGATCG	250
	TCGCCATCAA	CAAGATGGAC	ACGGCCAAGT	GGGCTGAGGC	TCGTTACCAG	300
	GAGATCATCA	AGGAGACCTC	CAACTTCATC	AAGAAGGTCTG	GCTACAACCC	350
	CAAGACTGTT	GCCTTCGTCC	CCATCTCGGG	CTTCCACGGC	GACAACATGC	400
	TTACTCCCTC	GACCAACTGC	CCCTGGTACA	AGGGCTGGGA	GAAGGAGGGC	450
50	AAGAGCGGCA	AGGTTACCGG	TAAGACTCTG	CTGGACGCCA	TTGACGCCGT	500
	CGAGCCCCCC	AAGCGCCCCA	CGGACAAGCC	CCTGCGTCTG	CCCCTCCAGG	550
	ATGTCTACAA	GATCGGCGGT	ATCGGCACTG	TCCCTGTCTG	CCGTATCGAG	600
	ACTGGTGTCC	TGAAGCCCGG	CATGGTCTGC	ACCTTTGCCC	CGTCCAACGT	650
	CACCACTGAA	GTCAAGTCCG	TCGAGATGCA	CCACGAGCAG	CTTGTTGAGG	700
55	GTGTTCCCGG	CGACAACGTC	GGCTTCAACG	TCAAGAACGT	CTCCGTCAAG	750
	GAGATCCGTC	GTGGCAACGT	TGCCGGTGAC	TCCAAGAACG	ACCCCCCCTC	800
	GGGCGCCGCC	ACCTTCAACG	CCCAGGTCAT	TGTCCTGAAC	CACCCCGGCC	850
	AGGTCGGCAA	CGGCTACGCC	CGGTTCTGG	ACTGCCACAC	CGCCACATT	900
	GCCTGCAAGT	TCACCGAGAT	CCTTGAGAAG	ATCGACCGCC	GTACCGGCAA	950
60	GTCGGTTGAG	AACAACCCCA	AGTTCATCAA	GTCGGGTGAC	GCCGCCATTG	1000

TCAAGCTGAC GCCCTCGAAG CCCATGTGCG TTGAGGCCTT CACTGACTAC 1050  
 CCCCCT 1056

5

## 2) INFORMATION FOR SEQ ID NO: 903

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1366 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Penicillium marneffe*  
 (B) STRAIN: ATCC 58950

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 903

CAAGGCTGAG	CGTGAGCGTG	GTATCACCAT	CGATATTGCT	CTCTGGAAGT	50
TCCAGACTGC	CAAGTACGAG	GTTACCGTCA	TTGACGCCCC	CGGTCACCGT	100
GATTTTCATCA	AGAACATGAT	CACTGGTACC	TCCCAGGCCG	ATTGCGCTAT	150
TCTCATCATT	GCCTCTGGTA	CTGGTGAATT	CGAGGCTGGT	ATCTCCAAGG	200
ATGGCCAGAC	TCGTGAGCAC	GCTCTTTTGG	CTTTCACCCT	CGGTGTCCGT	250
CAGCTCATTG	TTGCCCTCAA	CAAGATGGAC	ACTTGCAAGT	GGTCTCAGGG	300
TGAGTACTCG	TACCTGCGTT	TGGCCTTGAA	TATCTTACTA	ATGCACCATA	350
GATCGTTACA	ACGAAATTGT	CAAGGAGACT	TCCAACCTCA	TCAAGAAGGT	400
CGGATACAAC	CCTAAGAACG	TTCTTTTCGT	TCCTATCTCC	GGTTTCAACG	450
GTGACAACAT	GCTTGAGCCC	TCCCCCAACT	GCCCCCTGGT	CAAGGGTTGG	500
GAGAAGGAGA	CCAAGGCCGG	TAAGGTCACT	GGTAAGACCC	TCCTCGAGGC	550
CATCGACGCC	ATTGAGCCCC	CTACCCGTCC	CGCCAACAAG	GTTAGTCCCT	600
CCTCGACTAC	TCAAACCCTC	CTCATAAGTT	CAGATTACTG	ACTCGTTTAC	650
AGCCCCCTCCG	TCTTCCCCTC	CAGGATGTCT	ACAAGATCGG	TGGTATTGGA	700
ACGGTTCCCG	TCGGTCGTGT	TGAGACTGGT	ACCATCGTTC	CTGGTATGGT	750
TGCTAACCTTG	TAAGTCACTC	TCCTCGCTTA	TCCTACCTGA	AATCATCATG	800
TGCTAACCTTG	ACACTCAGCG	CTCCCGCCAA	CGTCACCACT	GAAGTCAAGA	850
GTGTTGAAAT	GCACCACCAG	CAGCTCACTG	CCGGTCAGCC	CGGTGACAAC	900
GTTGGTTTCA	ACGTGAAGAA	CGTCTCCGTC	AAGGAAATCC	GTCGTGGTAA	950
CGTTGCTGGT	GACAGCAAGA	ACGACCCCCC	TGCCGGTGCT	GCCTCCTTCA	1000
ACGCCCAGGT	CATCGTCCTC	AACCACCCCG	GTCAGGTCGG	TGCTGGTTAC	1050
GCCCCAGTCC	TCGATTGCCA	CACTGCCCCA	ATTGCTTGCA	AGTTCGCTGA	1100
GCTCCTCGAG	AAGATTGACC	GTCGTACCGG	AAAGTCTGTT	GAGGACCACC	1150
CCAAGTTCAT	CAAGTCCGGT	GACGCTGCCA	TCGTCAAGAT	GATTCTTCC	1200
AAGCCCATGT	GTGTTGAGGC	TTTCACCGAG	TACCCTCCTC	TCGGTCGTTT	1250
CGCCGTTCGC	GAGTAAGTTT	TATCTCCGTT	GTCTATTTTC	CATCCTTCCC	1300
TTCTCCTCCG	TCTTCCATAT	ATATTTTTC	AGTTATATGT	GACTAACCAC	1350
AAATCACGGG	AATAGC				1366

50

## 2) INFORMATION FOR SEQ ID NO: 904

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 841 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Curvularia lunata*

(B) STRAIN: ATCC 26425

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 904

```

10 CTATTATCGT TGTGCGCGCT TCCGACGGTC AAATGCCCCA GACTCGTGAG      50
   CATCTGCTGC TCGCCCGCCA GGTCGGTGTC CAGAAGATCG TTGTCTTCGT      100
   CAACAAGGTC GATGCTGTTG AGGACAAGGA GATGTTGGAG CTCGTGCGAGA      150
   TGGAGATGCG CGAATTGCTC AGCAGCTACG GCTTCGAGGG CGACGAGACT      200
   CCCATCATCA TGGGATCTGC CCTCTGCGCC ATTGAGGGCC GCGAACCTGA      250
   GATTGGTGTC AACCGAATTG ATGAGCTGCT CGAGGCCGTT GATACTTGGA      300
15 TCCCCACCCC TCAGCGTGAG ACCGACAAGC CTTTCCTCAT GGCCGTCGAG      350
   GATGTCTTCT CCATTGCTGG TCGTGGCACT GTCGTCTCTG GCCGTGTCGA      400
   GCGAGGTATC CTGAAGCGCG ATGCTGAAGT TGAGCTCGTC GGCAAGGGCA      450
   CCGCCCCCAT CAAGACCAAG GTTACCGATA TCGAGACCTT CAAGAAGTCC      500
   TGGCAGGAGT TCGCGCTGG TGACAACTCC GGTCTCCTTC TTCGTGGTGT      550
20 CAAGCGTGAT GAAGTCCGCC GTGGTATGGT CGTTTCCGTC CCTGGACAGG      600
   TCAAGGCGCA CAAGAAGTTC CTTGTCTCCA TGTACGTGTT GAGCAAGGAG      650
   GAAGGTGGTC GTCACACTGG CTTCCGTGAG AACTACAGGC CGCAAATGTT      700
   CATCCGCACT GCCGACGAGT CGTGTGCCCT GTACTGGCCA GAAGGCACCG      750
   AGGACGCCCC TGACAAGCTT GTTATGCCCG GTGACAACGT CGAGATGGTT      800
25 TGCGAGCTCC ATGCACCACA CGTCTTGAG CCTGGTCAAC G              841

```

2) INFORMATION FOR SEQ ID NO: 905

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 967 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Aspergillus niger*

(B) STRAIN: ATCC 9508

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 905

```

45 CGGTGCTATC ATTGTCGTCG CCGCCTCCGA CGGTCAGATG TACGTTAACC      50
   TTAAAAGAAT AACTCTCCTT CAGTATATAT GCTTACACTG GCGATCAACA      100
   GGCCCCAGAC TCGTGAGCAC TTGCTGCTTG CTCGTGAGGT CGGTGTCCAG      150
   AAGATCGTTG TCTTCGTCAA CAAGGTCGAT GCTATCGATG ACCCCGAGAT      200
   GCTGGAGCTC GTTGAGCTGG AAATGCGCGA GCTTCTCAGC ACCTACGGAT      250
50 TCGAGGGTGA GGAGACCCCC ATCGTCTTCG GCTCTGCTCT CTGCGCCATT      300
   GAGGACCGCC GCCCCGACAT CGGTACCGAG CGTATCGACG CTCTCCTCGA      350
   GGCCGTTGAC ACCTGGATCC CCACTCCCCA GCGTGACCTT GACAAGCCTT      400
   TCTTGATGTC CATTGAGGAA GTTTTCTCCA TCCCCGGTCG TGGTACCGTC      450
   GCCTCCGGCC GTGTGAGCGG TGGTCTCCTG AAGCGTGATA GCGAGGTGTA      500
55 GATCATCGGT ACCACCAACG AGGTCATCAA GACCAAGGTT ACCGACATTG      550
   AGACCTTCAA GAAGTCCTGC TCCGAGTCCC GCGCCGGTGA CAACTCCGGT      600
   CTCTGCTCC GTGGTGTCCG CCGTGAGGAT CTCCGCCGTG GTATGGTCAT      650
   TGCCGCTCCT GGCAGCGCCA AGGCCAACG CAAGTTCATG GTCTCCATGT      700
   ACGTCCTGAC CGAGGCTGAR GGTGGTCGCC GTACCGGTTT CGGTGTCCAG      750
60 TACCGTCCCC AGCTGTTTCT CCGCACTGCC GGTAAGTAAA ATTGCATTCT      800

```

ATTCCGCTAC	TAGGGAACCA	TCTCTAATTC	TATTTGCTAC	AGATGAGGCT	850
GCTGAGTTCA	GCTTCCCCGA	CGGAGACCAG	TCCCGCCGTA	TCATGCCCCG	900
TGACAAACGTC	GAGATGATCG	TCAAGACCCA	CCGCCCCGTC	GCCGCCGAGG	950
CCGGTCAGCG	CTTCAAC				967

5

## 2) INFORMATION FOR SEQ ID NO: 906

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 852 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 15 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Bipolaris hawaiiensis*
  - (B) STRAIN: ATCC 26067
- 20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 906

TGGTGCTATT	ATTGTTGTTG	CCGCTTCCGA	CGGTCAAATG	CCCCAGACTC	50
GTGAGCATCT	GCTGCTCGCC	CGTCAGGTCG	GTGTTTCAGAA	GATCGTTGTC	100
TTCGTTAACA	AGGTCGACGC	TGTCGAGGAC	AAGGAGATGT	TGGAGCTTGT	150
CGAGATGGAG	ATGCGCGAAC	TGCTCAGCAG	CTATGGCTTC	GAGGGCGACG	200
AGACCCCTAT	CATCATGGGT	TCTGCTCTCT	GCGCCATTGA	AGGCCGTCAA	250
CCCCGACATTG	GTGTCGAACG	AATTGACGAG	CTGCTCGAGG	CTGTTGATAC	300
30 TTGGATTCCC	ACCCCTCAGC	GTGAGACCGA	AAAGCCTTTC	CTCATGGCCG	350
TCGAGGATGT	CTTCTCCATT	GCTGGTCGTG	GTACCGTCGT	CTCTGGCCGT	400
GTCTGAACGAG	GTATCCTGAA	GCGCGATGCT	GAAGTTGAGC	TTGTGGGCAA	450
GGGCAGCGCA	CCCATCAAGA	CCAAGGTTAC	CGATATCGAG	ACCTTCAAGA	500
AGTCTTGCGA	GGAGTCCCGC	GCTGGTGACA	ACTCCGGTCT	CCTTCTTCGT	550
35 GGTGTTAAGC	GTGATGAAGT	CCGCCGTGGT	ATGGTCGTTT	CCGTCCCTGG	600
ACAGGTTAAG	GCGCACAAGA	AGTTCCTTGT	CTCCATGTAT	GTGCTGAGCA	650
AGGAGGAAGG	TGGCCGACAC	ACTGGCTTCG	TGAGAACTA	CAGGCCGCAA	700
ATGTTCATCC	GCACTGCCGA	CGAGTCGTGT	GCCCTGTACT	GGCCAGAAGG	750
CACCGAGGAT	GCCCACGACA	AGCTTGTCAT	GCCCGGTGAC	AACGTCGAGA	800
40 TGGTTTGCGA	GCTCCATGCA	CCACACGTCT	TGGAGACTGG	TCAGCGCTTC	850
AA					852

## 45 2) INFORMATION FOR SEQ ID NO: 907

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 966 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 50 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Aspergillus flavus*
  - (B) STRAIN: ATCC 26947
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 907
- 60



```

GGTGCTATTG TTGTCGTTGC TGCTTCGGAT GGTTCATGATG TATGGACAGG      50
CCCTTTGCTA CTGAATGGTT TCAAGATCTC GCGCTTACAC GTATTATAAT      100
AGGCCCCAGA CCCGGGAGCA CTTGCTGCTT GCCCGTCAGG TCGGTGTCCA      150
GAAGATCGTC GTTTTTGTCA ACAAGATTGA TGCCGTTGAG GACCCTGAGA      200
5  TGTTCGAGCT TGTCGAGTTG GAAATGCGCG AGCTCCTTAG CAGCTACGGC      250
TTCGAGGGCG AAGAGACTCC CATCATCTTC GGTTCGCTC TGTGTGCTTT      300
GGAGGACCGT CGCCCCGACA TTGGTGCCGA GCGTATCGAC GAGCTCATGA      350
AGGCCGTTGA CACCTGGATC CCTACCCCTC AGCGTGATCT TGACAAGCCT      400
10 TTCTCATGT CTGTCGAGGA AGTCTTCTCC ATCGCCGGTC GTGGTACCGT      450
TGCTCCGGC CGTGTCAAC GTGGTATCCT GAAGAAGGAC AGCGAAGTCG      500
AGATCATCGG AGGTAGCTTC GATGCTACCA AGACCAAGGT CACCGACATT      550
GAGACCTTCA AGAAGTCTTG TGACGAGTCC CGCGCTGGTG ACAACTCTGG      600
CTTGCTTCTC CGTGGTATCC GTCGTGAAGA CGTCCGCCGC GGAATGATCA      650
TTGCTGCTCC TGGCAGCACC AAGGCCACG ACCAGTTCTT GGTGTCCATG      700
15 TACGTTCTCA CTGAGGCTGA GGGTGGTCGT CGTACTGGCT TCGGCTCCAA      750
CTACCGCCCC CAGGTGTTCTG TTCGCACTGC TGGTAAGTCA AGCCTTTTGC      800
TCACTTAACG GTATTGATTA AGTTCTAACT GTTGATATCCT AGATGAGGCT      850
GCTGACCTCA GCTTCCCCGA CGGTGATGAG TCCCGGAGGG TGATGCCTGG      900
TGACAACGTC GAGATGGTCC TCAAGACTCA CCGCCCATTT GCTGCTGAGG      950
20 CTGGCCAGCG CTTCAA      966

```

## 2) INFORMATION FOR SEQ ID NO: 908

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 845 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 30 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Alternaria alternata*  
 35 (B) STRAIN: ATCC 62099

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 908

```

40 GGTGCTATCA TCGTCGTTGC TGCTTCCGAT GGTTCAGATGC CCCAGACCCG      50
TGAGCACTTG CTGCTCGCCC GTCAGGTCGG TGTTTCAAGAG ATCGTTGTCT      100
TCGTCAACAA GGTTCGATGCT GTCGAAGACC CGGAGATGTT GGAACCTCGTC      150
GAGATGGAGA TGCGTGAGTT ACTCACCAGC TACGGCTTCG AGGGCGACGA      200
GACACCCATC ATCATGGGTT CCGCTCTATG CGCCATCGAG GGCCGCCAGC      250
45 CCGAGATCGG TGTTACCAAG GTCGACGAGC TAATGGACGC TGTCGACTCA      300
TGGATCCCCA CCCCTCAGCG TGAGACCGAG AAGCCTTTCC TCATGGCTGT      350
TGAGGATGTC TTCTCGATTG CTGGACGTGG TACCGTCGTT TCGGGCCGTG      400
TCGAGCGCGG TATCTTGAAG CGTGACGCTG AAGTCGAGCT TGTCGGCAAG      450
GGCACC GCGC CAATCAAGAC CAAGGTCACT GATATTGAGA CCTTCAAGAA      500
50 GTCGTGCGAG GAGTCGCGCG CGGGTGATAA CTCCGGTCTT CTCCTCCGTG      550
GTGTCAAGCG TGATGACGTT CGCCGCGGTA TGTTTGTTC CGTTCCCGGA      600
CAAGTCAAGG CTCACAAGAA GTTCCTTGTC TCCATGTACG TTCTAAGCAA      650
AGAGGAGGGT GGTTCGTCACA CCGGCTTCGG CGAGAACTAC AGGCCGCAA      700
TGTTTCATCCG AACTGCCGAT GAATCCTGCG CACTTCACTT CCCAGAGGGT      750
55 ACCGAGGATG CGCACGACAA GCTAGTTATG CCCGGTGACA ACGTCGAGAT      800
GGTCTGCGAA CTCCACCAGC CCCACGTTCT AGAGACCGGT CAGCG      845

```

## 60 2) INFORMATION FOR SEQ ID NO: 909

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 931 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Penicillium marneffei*  
 (B) STRAIN: ATCC 64101

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 909

```

15  CGCTGTTGTT  GTCGTCGCTG  CTTCTGATGG  TCAAATGTAA  CATATCCACG      50
    AGCTGCCAAT  TATGGACACT  GCTGATAAGA  ATAGGCCCCA  AACCCGTGAG      100
    CACTTGCTCC  TCGCCCGTCA  GGTCGGTGTT  CAAAAGATCG  TCGTCTTCGT      150
    CAACAAGGTT  GATGCCGTCG  AGGACCCCGA  GATGTTGGAA  CTTGTCTGAAT      200
20  TGGAAATGCG  TGAAGTCTTG  ACCACCTACG  GTTTCGAGGG  TGAAGAGACC      250
    CCTATCATTT  TCGGATCCGC  TCTTTGCGCC  TTGGAAGGCC  GCAAGCCCGA      300
    GATTGGCGAA  CAGAAGATTG  ACGAGCTCAT  GAACGCCGTT  GATACCTGGA      350
    TCCCCACCCC  CCAGCGTGAC  CTTGACAAGC  CCTTCTTGAT  GTCCGTTGAG      400
    GAAGTTTTCT  CCATCTCTGG  TCGTGGTACC  GTTGCATCTG  GTCGTGTTGA      450
25  GCGTGGTATT  TTGCGCAAGG  ATTCTGAGGT  TGAGATTATC  GGATACCAGA      500
    AGAACCCTAT  CAAGACCAAG  GTTACCGACA  TTGAGACCTT  CAAGAAGTCT      550
    TGCGATGAAT  CTCGTGCTGG  TGACAACTCT  GGCTTGCTTC  TCCGTGGTAT      600
    CAAGCGTGAG  GACATTTCGT  GTGGTATGGT  TATCGCTGCT  CCTGGAACCA      650
    CCAAGGCTCA  TGACAACTTC  TTGGTCTCCA  TGTATGTCTT  GACTGAGGCT      700
30  GAAGGTGGTC  GTCGTACTGG  ATTCGGCGCC  AACTACCGTC  CTCAAGCTTT      750
    CATCCGTACT  GCCGGTATGT  TCCCTTTCAA  AGTCAATTAA  TGAGCGATTT      800
    GCTAACGAGT  TATAGATGAG  GCTGCTACTC  TCAGCTTCCC  CGGTGACGAT      850
    CAGTCCAAGC  AGGTCATGCC  CGGTGACAAC  GTTGAGATGA  TCTTGAAGAC      900
    ACACCGTCCC  GTTGCCGCCG  AAGCTGGTCA  G              931
35

```

## 2) INFORMATION FOR SEQ ID NO: 910

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 931 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Penicillium marneffei*  
 (B) STRAIN: ATCC 58950

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 910

```

55  CGCTGTTGTT  GTCGTCGCTG  CTTCTGATGG  TCAAATGTAA  CATATCCACG      50
    AGCTGCCAAT  TATGGACACT  GCTGATAAGA  ATAGGCCCCA  AACCCGTGAG      100
    CACTTGCTCC  TCGCCCGTCA  GGTCGGTGTT  CAAAAGATCG  TCGTCTTCGT      150
    CAACAAGGTT  GATGCCGTCG  AGGACCCCGA  GATGTTGGAA  CTTGTCTGAAT      200
    TGGAAATGCG  TGAAGTCTTG  ACCACCTACG  GTTTCGAGGG  TGAAGAGACC      250
    CCTATCATTT  TCGGATCCGC  TCTTTGCGCC  TTGGAAGGCC  GCAAGCCCGA      300
60  GATTGGCGAA  CAGAAGATTG  ACGAGCTCAT  GAACGCCGTT  GATACCTGGA      350

```

	TCCCCACCCC	CCAGCGTGAC	CTTGACAAGC	CCTTCTTGAT	GTCCGTTGAG	400
	GAAGTTTTCT	CCATCTCTGG	TCGTGGTACC	GTTGCATCTG	GTCGTGTTGA	450
	GCGTGGTATT	TTGCGCAAGG	ATTCTGAGGT	TGAGATTATC	GGATACCAGA	500
	AGAACCCTAT	CAAGACCAAG	GTTACCGACA	TTGAGACCTT	CAAGAAGTCT	550
5	TGCGATGAAT	CTCGTGCTGG	TGACAACTCT	GGCTTGCTTC	TCCGTGGTAT	600
	CAAGCGTGAG	GACATTTCGT	GTGGTATGGT	TATCGCTGCT	CCTGGAACCA	650
	CCAAGGCTCA	TGACAACTTC	TTGGTCTCCA	TGTATGTCTT	GACTGAGGCT	700
	GAAGGTGGTC	GTCGTACTGG	ATTCGGCGCC	AACTACCGTC	CTCAAGCTTT	750
	CATCCGTAAT	GCCGGTATGT	TCCCTTTCAA	AGTCAATTAA	TGAGCGATTT	800
10	GCTAACGAGT	TATAGATGAG	GCTGCTACTC	TCAGCTTCCC	CGGTGACGAT	850
	CAGTCCAAGC	AGGTCATGCC	CGGTGACAAC	GTTGAGATGA	TCTTGAAGAC	900
	ACACCGTCCC	GTTGCCGCCG	AAGCTGGTCA	G		931

15

## 2) INFORMATION FOR SEQ ID NO: 911

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 911

GACGGMKKCA TGCCGCARAC

20

30

## 2) INFORMATION FOR SEQ ID NO: 912

## (i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 912

GACGGCGKCA TGCCGCARAC

20

45

## 2) INFORMATION FOR SEQ ID NO: 913

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

55

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 913

60 GACGGYSYCA TGCKCAGAC

20

## 2) INFORMATION FOR SEQ ID NO: 914

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 914

GAARAGCTGC GGRCGRTAGT G

21

## 2) INFORMATION FOR SEQ ID NO: 915

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 915

AAACGCCTGA GGRCGGTAGT T

21

## 2) INFORMATION FOR SEQ ID NO: 916

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 916

GCCGAGCTGG CCGGCTTCAG

20

## 2) INFORMATION FOR SEQ ID NO: 917

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 917

TCGTGCTACC CGTYGCCGCC AT

22

5

2) INFORMATION FOR SEQ ID NO: 918

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1391 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: *Escherichia coli*  
 (E) ACCESSION NUMBER: J01672

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 918

25	AGAGAAGCCT	GTCGGCACCG	TCTGGTTTGC	TTTTGCCACT	GCCCGCGGTG	50
	AAGGCATTAC	CCGGCGGGAT	GCTTCAGCGG	CGACCGTGAT	GCGGTGCGTC	100
	GTCAGGCTAC	TGCGTATGCA	TTGCAGACCT	TGTGGCAACA	ATTTCTACAA	150
	AACACTTGAT	ACTGTATGAG	CATACAGTAT	AATTGCTTCA	ACAGAACATA	200
	TTGACTATCC	GGTATTACCC	GGCATGACAG	GAGTAAAAAT	GGCTATCGAC	250
	GAAAACAAAC	AGAAAGCGTT	GGCGGCAGCA	CTGGGCCAGA	TTGAGAAACA	300
30	ATTTGGTAAA	GGCTCCATCA	TGCGCCTGGG	TGAAGACCGT	TCCATGGATG	350
	TGGAAACCAT	CTCTACCGGT	TGCGCTTTCAC	TGGATATCGC	GCTTGGGGCA	400
	GGTGGTCTGC	CGATGGGCCG	TATCGTCGAA	ATCTACGGAC	CGGAATCTTC	450
	CGGTAAAACC	ACGCTGACGC	TGCAGGTGAT	CGCCGCAGCG	CAGCGTGAAG	500
	GTAAAACCTG	TGCGTTTATC	GATGCTGAAC	ACGCGCTGGA	CCCAATCTAC	550
35	GCACGTAAAC	TGGGCGTCGA	TATCGACAAC	CTGCTGTGCT	CCCAGCCGGA	600
	CACCGGCGAG	CAGGCACTGG	AAATCTGTGA	CGCCCTGGCG	CGTTCTGGCG	650
	CAGTAGACGT	TATCGTCGTT	GACTCCGTGG	CGGCACTGAC	GCCGAAAGCG	700
	GAAATCGAAG	GCGAAATCGG	CGACTCTCAC	ATGGGCCTTG	CGGCACGTAT	750
	GATGAGCCAG	GCGATGCGTA	AGCTGGCGGG	TAACCTGAAG	CAGTCCAACA	800
40	CGCTGCTGAT	CTTCATCAAC	CAGATCCGTA	TGAAAATTGG	TGTGATGTTC	850
	GGTAACCCGG	AAACCACTAC	CGGTGGTAAC	GCGCTGAAAT	TCTACGCCTC	900
	TGTTTCGTCTC	GACATCCGTC	GTATCGGCGC	GGTGAAAGAG	GGCGAAAACG	950
	TGGTGGGTAG	CGAAACCCGC	GTGAAAGTGG	TGAAGAACAA	AATCGCTGCG	1000
	CCGTTTAAAC	AGGCTGAATT	CCAGATCCTC	TACGGCGAAG	GTATCAACTT	1050
45	CTACGGCGAA	CTGGTTGACC	TGGGCGTAAA	AGAGAAGCTG	ATCGAGAAAG	1100
	CAGGCGCGTG	GTACAGCTAC	AAAGGTGAGA	AGATCGGTCA	GGGTAAAGCG	1150
	AATGCGACTG	CCTGGCTGAA	AGATAACCCG	GAAACCGCGA	AAGAGATCGA	1200
	GAAGAAAGTA	CGTGAGTTGC	TGCTGAGCAA	CCCGAACTCA	ACGCCGGATT	1250
	TCTCTGTAGA	TGATAGCGAA	GGCGTAGCAG	AAACTAACGA	AGATTTTTAA	1300
50	TCGTCTTGTT	TGATACACAA	GGGTCGCATC	TGCGGCCCTT	TTGCTTTTTT	1350
	AAGTTGTAAG	GATATGCCAT	GACAGAATCA	ACATCCCGTC	G	1391

55 2) INFORMATION FOR SEQ ID NO: 919

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single

60

718

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 919

GGICCGART CITMIGGIAA RAC

23

10

2) INFORMATION FOR SEQ ID NO: 920

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

15 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 920

TCICCVATIT CICCITCIAI YTC

23

25

2) INFORMATION FOR SEQ ID NO: 921

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

30 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 921

TIYRTIGAYG CIGARCAIGC

20

40

2) INFORMATION FOR SEQ ID NO: 922

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 922

55 TARAAYTTIA RIGCIYKICC ICC

23

60

2) INFORMATION FOR SEQ ID NO: 923

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 923

GACGCIGCCA TCCTGATGAT C

21

2) INFORMATION FOR SEQ ID NO: 924

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 924

ACCTCAGTCG TCACGTTGGC G

21

2) INFORMATION FOR SEQ ID NO: 925

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 925

AAGCAGATGG TTGTGTGCTG

20

2) INFORMATION FOR SEQ ID NO: 926

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 926

CAGCTGCTCG TGGTGCATCT CGAT

24

## 2) INFORMATION FOR SEQ ID NO: 927

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 19 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 927

15 ACGCGGAGAA GGTGCGCTT 19

## 2) INFORMATION FOR SEQ ID NO: 928

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

25 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 928

30 GGTCGTTCTT CGAGTCACCG CA 22

## 2) INFORMATION FOR SEQ ID NO: 929

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 448 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

45 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Bacteroides fragilis*  
 (B) STRAIN: ATCC 25285

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 929

50 TTCAGCATGC CATTTCAAAA CAGGCCGAAG CCGATATCGT GATTATCGCT 50  
 GCTTGTGGGG AGCGTGCAAA TGAAGTTGTG GAAATCTTTA CCGAATTTCC 100  
 GGAATTGGTG GACCCGCACA CGGGACGTAA GCTGATGGAG CGTACCATTA 150  
 TTATCGCAAA TACATCGAAC ATGCCGGTAG CAGCGCGTGA AGCTTCTGTG 200  
 55 TATACGGCCA TGACGATTGC CGAATACTAT CGTGCCATGG GATTGAAAGT 250  
 CCTGCTGATG GCAGACTCCA CTTCCCGTTG GGCGCAGGCA TTGCGTGAGA 300  
 TGTCGAACCG TATGGAGGAG TTGCCCGGAC CGGATGCATT CCCGATGGAC 350  
 CTGTCCTCAA TCATTCTTAA CTTCTATGGC CGTGCAGGCT ACGTGAAACT 400  
 60 GAATAACGGC GAGAGCGGTT CTATTACCTT TATCGGTACA GTATCACC 448



## 2) INFORMATION FOR SEQ ID NO: 930

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 438 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 10 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Bacteroides distasonis*  
 15 (B) STRAIN: ATCC 8503

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 930

	GCTATCTCTA	AACAAGCGGA	AGCGGATATC	GTGATTATCG	CCGCCTGCGG	50
20	TGAGCGTGCG	AATGAGGTCG	TAGAGGTATT	TACGGAGTTC	CCGGAATTGG	100
	TAGACCCGCA	TACGGGACGT	AAATTGATGG	AACGTACGAT	CATTATCGCC	150
	AATACATCCA	ACATGCCGGT	AGCCGCTCGT	GAGGCATCCG	TATATACGGC	200
	GATGACCATC	GCCGAGTATT	ATCGCAGCAT	GGGTTTGAAG	GTTCTGTTGA	250
	TGGCCGACTC	TACTTCCCGC	TGGGCACAGG	CTTTGCGTGA	GATGTCCAAC	300
25	CGTTTGGAGG	AGTTGCCGGG	ACCGGATGCT	TTCCCGATGG	ACTTGTCCGC	350
	TATCGTGGCG	AACTTCTACG	CTCGTGCGGG	ATTTCGTTTCAT	TTGAATAACA	400
	ACGCTACAGG	CTCCGTCACT	TTCATCGGTA	CGGTATCG		438

## 30 2) INFORMATION FOR SEQ ID NO: 931

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 453 bases  
 35 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- 40 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Porphyromonas asaccharolytica*  
 (B) STRAIN: ATCC 25260

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 931

	CCTCCAGCAC	GCTATCTCTA	AGCAGGCGGA	GGCTGATATC	GTCATTATGG	50
	CAGCCTGCGG	TGAGCGTGCT	AATGAGGTGG	TGGAGATCTT	TGCCGAGTTC	100
	CCTGAGCTCG	AAGACCCACA	CACGGGACGC	AAGCTGATGG	AGCGTACGAT	150
50	CATCATCGCT	AACACGAGTA	ACATGCCAGT	GGCTGCTCGT	GAGGCTTCGG	200
	TCTACACCGC	TATGACCATC	GCTGAGTACT	ACCGCTCGAT	GGGTCTCAAA	250
	GTA CTCTAA	TGGCTGACTC	GACCTCTCGC	TGGGCACAGG	CACTGCGTGA	300
	GATGTCTAAC	CGTCTAGAGG	AGCTGCCTGG	ACCAGATGCA	TTCCCGATGG	350
	ACTTGTCCGC	TATCGTGGCA	AACTTCTACG	CTCGTGCCGG	CTTCGTCTAT	400
55	CTCAACAACG	GTGAGACAGG	TTCTGTAACC	TTCATCGGTA	CGGTCTCTCC	450
	AGC					453

## 60 2) INFORMATION FOR SEQ ID NO: 932

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 835 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Listeria monocytogenes*  
 (B) STRAIN: ATCC 15313

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 932

```

CGGAGCTATC TTAGTAGTAT CTGCTGCTGA TGGCCCAATG CCACAAACTC      50
GTGAACATAT CTTACTTTCA CGTCAAGTTG GTGTTCCATA CATCGTTGTA      100
TTCATGAACA AATGTGACAT GGTGACGAT GAAGAATTAC TAGAATTAGT      150
TGAAATGGAA ATTCGTGATC TATTAAGTGA ATATGAATTC CCTGGCGATG      200
ACATTCCCTGT AATCAAAGGT TCAGCTCTTA AAGCACTTCA AGGTGAAGCT      250
GACTGGGAAG CTAAAATTGA CGAGTTAATG GAAGCTGTAG ATTCTTACAT      300
TCCAACTCCW GAACGTGATA CTGACAAACC ATTCATGATG CCAGTTGAGG      350
ATGTATTCTC AATCACTGGT CGTGGAACAG TTGCAACTGG ACGTGGTTGAA      400
CGTGGACAAG TTAAAGTTGG TGACGAAGTA GAAGTTATCG GTATCGAAGA      450
AGAAAGCAAA AAAGTAGTAG TAACTGGAGT AGAAATGTTC CGTAAATTAC      500
TAGACTACGC TGAAGCTGGC GACAACATTG GCGCACTTCT ACGTGGTGTT      550
GCTCGTGAAG ATATCCAACR TGGTCAAGTA TTAGCTAAAC CAGGTTTCGAT      600
TACTCCACAC ACTAATTCA AAGCTGAAAC TTATGTTTTA ACTAAAGAAG      650
AAGGTGGACG TCACACTCCA TTCTTCAACA ACTACGCCC ACAATTCTAT      700
TTCCGTACTA CTGACGTAAC TGGTATTGTT ACACTTCCAG AAGGTACTGA      750
AATGGTAAYG CCTGGTGATA ACATTGAGCT TGCAGTTGAA CTAATTGCAC      800
CAATCGCTAT CGAAGACGGT ACTAAATTCT CTATC                      835

```

## 2) INFORMATION FOR SEQ ID NO: 933

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 933

CATCATCGTI TTCMTGAACA ARTG

24

## 2) INFORMATION FOR SEQ ID NO: 934

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 934

TCACGYTTRR TACCACGCAG IAGA

24

5

2) INFORMATION FOR SEQ ID NO: 935

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
15 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 935

20 GGIAARWSIC ARYTITGYCA YAC

23

2) INFORMATION FOR SEQ ID NO: 936

25 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
30 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 936

35 TCISIIYTCIG GIARRCAIGG

20

40 2) INFORMATION FOR SEQ ID NO: 937

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
45 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 937

ATACIGARG YITTYGGIGA RTT

23

55 2) INFORMATION FOR SEQ ID NO: 938

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
60 (C) STRANDEDNESS: Single

724

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 938

CYIGTIGYIS WIGCRTGIGC

20

10

2) INFORMATION FOR SEQ ID NO: 939

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 1203 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*  
 (C) ACCESSION NUMBER: D10023

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 939

	ATGTCTCAAG	TTCAAGAACA	ACATATATCA	GAGTCACAGC	TTCAGTACGG	50
	GAACGGTTCG	TTGATGTCCA	CTGTACCAGC	AGACCTTTCA	CAGTCAGTCG	100
	TTGATGGAAA	CGGCAACGGT	AGCAGCGAAG	ATATTGAGGC	CACCAACGGC	150
30	TCCGGCGATG	GTGGCGGATT	GCAGGAGCAA	GCGGAAGCGC	AAGGTGAAAT	200
	GGAGGATGAA	GCATACGATG	AAGCTGCCTT	AGGTTTCGTTT	GTGCCAATAG	250
	AAAAACTGCA	AGTGAACGGG	ATTACTATGG	CGGATGTGAA	AAAACATAAGG	300
	GAGAGTGGGC	TTCACACTGC	TGAAGCGGTA	GCATATGCTC	CCAGAAAGGA	350
	TTTATTGGAA	ATCAAAGGTA	TATCGGAAGC	TAAGGCAGAT	AAGTTGCTAA	400
35	ACGAAGCGGC	AAGGCTAGTG	CCTATGGGAT	TTGTCACGGC	TGCTGATTTT	450
	CATATGAGAA	GATCGGAGCT	GATTTGTTTG	ACAACGGGTT	CTAAGAATTT	500
	GGACACTCTT	TTGGGTGGTG	GTGTGGAAAC	TGGTTCTATT	ACTGAGCTTT	550
	TCGGTGAATT	CAGGACAGGT	AAGTCCCAGC	TATGTCACAC	TTTGGCCGTG	600
	ACATGCCAAA	TTCCATTGGA	TATTGGTGGC	GGTGAAGGTA	AGTGTTTGTA	650
40	TATCGATACC	GAAGGTACTT	TCAGGCCGGT	AAGATTGGTA	TCCATAGCTC	700
	AGCGGTTCGG	ATTAGACCCG	GATGATGCTT	TGAACAACGT	TGCGTATGCA	750
	AGAGCCTATA	ACGCCGATCA	TCAGTTAAGA	CTTCTGGATG	CTGCTGCCCA	800
	AATGATGAGC	GAGTCTCGGT	TTTCCTTGAT	TGTGGTCGAT	TCTGTTATGG	850
	CTCTATACCG	TACGGATTTT	TCTGGTCGTG	GTGAACTAAG	CGCAAGGCAA	900
45	ATGCATTTAG	CCAAATTTAT	GCGTGCTTTG	CAAAGGCTGG	CCGACCAATT	950
	TGGTGTGTC	GTCGTCGTTA	CTAACCAAGT	GGTCGCCCAA	GTTGATGGTG	1000
	GTATGGCTTT	TAATCCAGAT	CCAAAGAAGC	CTATCGGTGG	TAATATTATG	1050
	GCACATTCTT	CCACCACGCG	ATTAGGTTTC	AAAAAGGGTA	AGGGATGTCA	1100
	AAGATTATGC	AAAGTTGTTG	ACTCACCTTG	CTTACCAGAG	GCTGAATGTG	1150
50	TGTTTCGCGAT	CTATGAAGAT	GGTGTGGTG	ACCCACAGAGA	AGAAGACGAG	1200
	TAG					1203

55 2) INFORMATION FOR SEQ ID NO: 940

(i) SEQUENCE CHARACTERISTICS:

60 (A) LENGTH: 1800 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Saccharomyces cerevisiae*

(B) STRAIN: GRF88

(C) ACCESSION NUMBER: M87549

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 940

	CGATCCAATT	GCTGGTCTTA	AGATGCATTT	GATTGATCTA	GGTATTGCCA	50
	CTGAAGCTGA	AGTCAAAGCT	TACGACAAGT	CCGCTAGAAA	ATACGTTGAC	100
	GAACAAGTTG	AATTAGCTGA	TGCTGCTCCT	CCTCCAGAAG	CCAAATTATC	150
15	CATCTTGTTT	GAAGACGTCT	ACGTGAAAGG	TACAGAAACT	CCAACCCCTAA	200
	GAGGTAGGAT	CCCTGAAGAT	ACTTGGGACT	TCAAAAAGCA	AGGTTTTGCC	250
	TCTAGGGATT	AATTAAATCG	TAAGGAAAAA	TAAAATAATA	GTGCTGTGAT	300
	CGCATGATAT	TCTTCCCTGG	AAGCGCCATT	TTATAGCAAG	AAATGTAAGT	350
	CAAGTATATT	TTAACTGTAT	ATACAACAAT	ATGACTCTTT	TTTATGCCCTT	400
20	GTTGTTTTTC	TTCGGGTTTT	CCCACACATT	GTGTGGAGAG	ATAGTTATTA	450
	ACAGACCGAA	AATAGCCGCC	CAAGGATAAA	CTTTTATATA	AAGGGAAGGG	500
	TAGTTGACCC	AAAAATTTGG	ATTCTACTTT	CCAGATTTAC	TTTCACCCCTT	550
	TTATATTTGC	TGTAGTCTGT	TATGCCAATG	AGGAAAGCAT	TTGAACAAAT	600
	ATGTCTGTTA	CAGGAAGTGA	GATCGATAGT	GATACAGCAA	AAAATATTCT	650
25	TAGTGTAGAT	GAATTACAGA	ACTATGGTAT	TAATGCCTCA	GATCTTCAAA	700
	AATTGAAGTC	TGGTGGGATA	TACACAGTCA	ATGTATGTTA	TAATAACATT	750
	TTTAAACCTT	CTGCTGTAGA	GGTTCCTTCC	CCCTTTCTTT	TACTAACTAA	800
	TAATTTGGAA	AGGAACTTTT	ATAGACCGTT	TTGTCAACAA	CAAGAAGACA	850
	TCTATGTAAA	ATTAAAGGGT	TAAGTGAGGT	GAAAGTGGAA	AAAATTAAAG	900
30	AAGCTGCTGG	AAAAATCATA	CAGGTGGGAT	TTATCCCCGC	TACTGTACAA	950
	TTAGATATAA	GACAGCGTGT	GTACTCCTTA	TCAACTGGAT	CTAAGCAACT	1000
	AGATTCAATC	CTAGGTGGTG	GAATAATGAC	AATGAGTATC	ACTGAAGTAT	1050
	TTGGTGAAAT	TAGGTGTGGT	AAGACACAGA	TGTCTCATA	TTTGTGTGTT	1100
	ACCACGCAGC	TTCCGAGGGA	AATGGGTGGT	GGTGAAGGGA	AAGTAGCATA	1150
35	TATTGATACA	GAAGGCACTT	TCAGGCCCGA	GAGGATTAAG	CAAATTGCAG	1200
	AAGGTTATGA	ATTGGATCCC	GAGTCATGTT	TGGCAAACGT	TTCATATGCT	1250
	AGAGCCTTGA	ATAGTGAACA	TCAAATGGAA	CTTGTTGAAC	AATTGGGTGA	1300
	AGAACTTAGT	TCTGGAGATT	ATCGCCTTAT	CGTGGTAGAT	TCTATAATGG	1350
	CAAACCTCAG	AGTAGACTAC	TGCGGTAGAG	GTGAACTAAG	CGAAAGACAG	1400
40	CAAAAGCTAA	ATCAACATCT	TTTCAAATTG	AATAGATTGG	CAGAGGAATT	1450
	TAATGTTGCA	GTATTTCTGA	CAAACCAAGT	TCAATCAGAC	CCAGGTGCTT	1500
	CTGCATTATT	TGCCTCGGCA	GATGGTAGGA	AACCAATTGG	AGGGCACGTT	1550
	CTGGCACATG	CGTCAGCAAC	AAGGATTTTG	TTGAGAAAAG	GGCGTGGTGA	1600
	CGAAAGAGTT	GCCAAGTTAC	AAGATTCCCC	AGATATGCCT	GAAAAAGAAT	1650
45	GTGTCTACGT	AATTGGTGAA	AAAGGTATTA	CCGATTCAAG	TGACTAGTTT	1700
	TTGTATACTT	TTTTAATGAA	GATGACATTG	CTCCTTTATT	AAACTTTTCT	1750
	TTTACTTTGT	GTTACTAATA	TTATTAATAT	CTTGTTATGA	TTCTTTGTTT	1800

50

## 2) INFORMATION FOR SEQ ID NO: 941

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 430 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Cryptococcus humicola*  
 (B) STRAIN: ATCC 38294

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 941

```

CGTCCTTATC CAGGAGCTCA TTAACAACAT TGCCAAGGCC CACGGTGGTT      50
TCTCCGTCTT CACCGGTGTC GGTGAGCGTA CCCGTGAGGG TAACGACCTG      100
TACCACGAGA TGCGTGAGAC TGGTGTGATC AACCTCGAGG GCGACTCCAA      150
GGTTCGCTCTC GTCTTCGGCC AGATGAACGA GCGGCGGCGA GCGGCTGCCC      200
GTGTGCGCCCT TACCGGCCTC ACCATCGCCG AGTACTTCCG TGACGAGGAG      250
GGTCAGGACG TGCCTTCTCTT CATCGACAAC ATTTTCCGTT TCACCCAGGC      300
CGGTTCCGAG GTGTCTGCCC TTCTCGGTCG TATCCCCTCG GCCGTCGGTT      350
ACCAGCCAC CCTCGCTACC GACATGGGTT CCATGCAGGA GCGTATCACC      400
ACCACCAAGA AGGGTTCGAT TACCTCCGTC      430

```

## 2) INFORMATION FOR SEQ ID NO: 942

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 794 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 43895

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 942

```

CGTGACGAT GCTCTTGAGG TGCAAAATGG TAATGAGCGT CTGGTGCTGG      50
AAGTTCAGCA GCAGCTCGGC GCGGCTATCG TGCGTACCAT CGCAATGGGT      100
TCCTCCGACG GTCTGCGTCG CGGTCTGGAT GTAAAAGACC TCGAACACCC      150
GATCGAAGTC CCGGTAGGTA AAGCGACTCT GGGCCGTATC ATGAACGTAC      200
TGGGTGAACC GGTGACATG AAAGGCGAGA TCGGTGAAGA AGAGCGTTGG      250
GCGATTACAC GCGCAGCACC TTCCTACGAA GAGCTGTCAA ACTCTCAGGA      300
ACTGCTGGAA ACCGGTATCA AAGTTATCGA CCTGATGTGT CCGTTCGCTA      350
AGGGCGGTAA AGTTGGTCTG TTCGGTGGTG CGGGTGTAGG TAAAACCGTA      400
AACATGATGG AGCTCATTCG TAACATCGCG ATCGAGCACT CCGGTTACTC      450
TGTGTTTTCG GCGGTAGGTG AACGTACTCG TGAGGGGAAC GACTTCTACC      500
ACGAAATGAC CGACTCCAAC GTTATCGATA AAGTATCCCT GGTGTATGGC      550
CAGATGAACG AGCCGCCGGG AAACCGTCTG CGCGTAGCTC TGACCGGTCT      600
GACCATGGCT GAGAAATTCG GTGACGAAGG TCGTGACGTT CTGCTGTTCT      650
TTGACAACAT CTATCGTTAC ACCCTGGCCG GTACGGAAGT ATCCGCACTG      700
CTGGGCCCGTA TGCCTTCAGC GGTAGGTTAT CAGCCGACCC TGGCGGAAGA      750
GATGGGCGTT CTGCAGGAAC GTATCACCTC CACCAAAACC GGTT          794

```

## 2) INFORMATION FOR SEQ ID NO: 943

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 814 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 35401

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 943

20	ATGCCGTACC	GCGCGTGTAC	GATGCTCTTG	AGGTGCAAAA	TGGTAATGAG	50
	CGTCTGGTGC	TGGAAGTTCA	GCAGCAGCTC	GGCGGCGGTA	TCGTACGTAC	100
	CATCGCAATG	GGTTCTCTCCG	ACGGTCTGCG	TCGCGGTCTG	GATGTAAAAG	150
	ACCTCGAACA	CCCGATTGAA	GTCCCGGTAG	GTAAAGCGAC	TCTGGGCCGT	200
	ATCATGAACG	TACTGGGTGA	ACCGGTCGAC	ATGAAAGGCG	AGATCGGTGA	250
25	AGAAGAGCGT	TGGGCGATTTC	ACCGCGCAGC	ACCTTCCTAC	GAAGAGCTGT	300
	CAAACCTCTCA	GGAACGTCTG	GAAACCGGTA	TCAAAGTTAT	CGACCTGATG	350
	TGTCCGTTTC	CTAAGGGCGG	TAAAGTTGGT	CTGTTCCGGT	GTGCGGGTGT	400
	AGGTAAAACC	GTAAACATGA	TGGAGCTCAT	TCGTAACATC	GCGATCGAGC	450
	ACTCCGGTTA	CTCTGTGTTT	GCGGGCGTAG	GTGAACGTAC	TCGTGAGGGT	500
30	AACGACTTCT	ACCACGAAAT	GACCGACTCC	AACGTTATCG	ACAAAGTATC	550
	CCTGGTGTAT	GGCCAGATGA	ACGAGCCGCC	GGGAAACCGT	CTGCGCGTTG	600
	CTCTGACCGG	TCTGACCATG	GCTGAGAAAT	TCCGTGACGA	AGGTCGTGAC	650
	GTTCTGCTGT	TCGTTGACAA	CATCTATCGT	TACACCCTGG	CCGGTACGGA	700
	AGTATCCGCA	CTGCTGGGCC	GTATGCCTTC	AGCGGTAGGT	TATCAGCCGA	750
35	CCCTGGCGGA	AGAGATGGGC	GTTCTGCAGG	AACGTATCAC	CTCCACCAA	800
	ACTGGTTCTA	TCAC				814

## 2) INFORMATION FOR SEQ ID NO: 944

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 798 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 11775

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 944

55	GTGTACGATG	CTCTTGAGGT	GCAAAATGGT	AATGAGCGTC	TGGTGCTGGA	50
	AGTTTCAGCAG	CAGCTCGGCG	GCGGTATCGT	GCGTACCATC	GCAATGGGTT	100
	CCTCCGACGG	TCTGCGTCGC	GGTCTGGATG	TAAAAGACCT	CGAACACCCG	150
	ATCGAAGTCC	CGGTAGGTAA	AGCGACTCTG	GGCCGTATCA	TGAACGTACT	200
60	GGGTGAACCG	GTCGACATGA	AAGGCGAGAT	CGGTGAAGAA	GAGCGTTGGG	250

	CGATTCACCG	CGCAGCACCT	TCCTACGAAG	AGCTGTCAAA	CTCTCAGGAA	300
	CTGCTGGAAA	CCGGTATCAA	AGTTATCGAC	CTGATGTGTC	CGTTCGCTAA	350
	GGGCGGTAAA	GTTGGTCTGT	TCGGTGGTGC	GGGTGTAGGT	AAAACCGTAA	400
	ACATGATGGA	GCTTATTCGT	AACATCGCGA	TCGAGCACTC	CGGTTACTCT	450
5	GTGTTTGC	GCGTAGGTGA	ACGTACTCGT	GAGGGTAACG	ACTTCTACCA	500
	CGAAATGACC	GACTCCAACG	TTATCGACAA	AGTATCCCTG	GTGTATGGCC	550
	AGATGAACGA	GCCGCCGGA	AACCGTCTGC	GCGTTGCTCT	GACCGGTCTG	600
	ACCATGGCTG	AGAAATTCCG	TGACGAAGGT	CGTGACGTTT	TGCTGTTCGT	650
	TGACAACATC	TATCGTTACA	CCCTGGCCGG	TACGGAAGTA	TCCGCACTGC	700
10	TGGGCCGTAT	GCCTTCAGCG	GTAAGTTATC	AGCCGACCCT	GGCGGAAGAG	750
	ATGGGCGTTC	TGCAGGAACG	TATCACCTCC	ACCAAACCG	GTTCTATC	798

## 15 2) INFORMATION FOR SEQ ID NO: 945

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 812 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 25922

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 945

30	ATGCCGTACC	GCGCGTGATC	GATGCTCTTG	AGGTGCAAAA	TGGTAATGAG	50
	CGTCTGGTGC	TGGAAGTTCA	GCAGCAGCTC	GGCGGCGGTA	TCGTGCGTAC	100
	CATCGCAATG	GGTTCCCTCCG	ACGGTCTGCG	TCGCGGTCTG	GATGTAAAAG	150
	ACCTCGAACA	CCCAGATCGAA	GTCCCGGTAG	GTAAAGCGAC	TCTGGGCCGT	200
35	ATCATGAACG	TACTGGGTGA	ACCGGTCGAC	ATGAAAGGCG	AGATCGGTGA	250
	AGAAGAGCGT	TGGGCGATTC	ACCGCGCAGC	ACCTTCCTAC	GAAGAGCTGT	300
	CAAACCTCTA	GGAACCTGCTG	GAAACCGGTA	TCAAAGTTAT	CGACCTGATG	350
	TGTCCGTTCT	CTAAGGGCGG	TAAAGTTGGT	CTGTTTCGGTG	GTGCGGGTGT	400
	AGGTAAAACC	GTAAACATGA	TGGAGCTTAT	TCGTAACATC	GCGATCGAGC	450
40	ACTCCGGTTA	CTCTGTGTTT	GCGGGCGTAG	GTGAACGTAC	TCGTGAGGGT	500
	AACGACTTCT	ACCACGAAAT	GACCGACTCC	AACGTTATCG	ACAAAGTATC	550
	CTTGGTGTAT	GGCCAGATGA	ACGAGCCGCC	GGGAAACCGT	CTGCGCGTTG	600
	CTCTGACCGG	TCTGACCATG	GCTGAGAAAT	TCCGTGACGA	AGGTCGTGAC	650
	GTATTGCTGT	TCGTGCGATA	CATCTATCGT	TACACCCTGG	CCGGTACCGA	700
45	AGTATCCGCA	CTGCTGGGCC	GTATGCCTTC	AGCGGTAGGT	TATCAGCCGA	750
	CCCTGGCGGA	AGAGATGGGC	GTTCTGCAGG	AACGTATCAC	CTCCACCAAA	800
	ACCGGTTCTA	TC				812

## 50 2) INFORMATION FOR SEQ ID NO: 946

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 832 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA



## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria polysaccharea*  
 (B) STRAIN: ATCC 43768

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 946

```

GCGACGCTAT CCCGCATGTT TACGATGCCC TGAAATTGGA CGAGAACGGT      50
CTGACTCTGG AAGTTCAACA ACTTCTGGGT GACGGCGTTG TCCGTACTAT      100
TGCAATGGGT AGTTCAGACG GCCTGAAACG CGGCATGTCT GTAAGCAATA      150
10 CTGGTGCGCC AATCACTGTG CCGGTAGGTA AAGGTACTTT GGGTCGTATT      200
GTCGACGTAT TGGGTACGCC TGTGTATGAA GCAGGTCCGA TCGATACCGA      250
CAAGAGCCGT GCCATTCAAC AAAC TGCTCC GAAATTCGAC GAGTTGTCTG      300
CAACTACCGA ATTGTTGGAA ACCGGTATTA AAGTGATCGA CTTGCTGTGT      350
CCGTTTGTCTA AAGGCGGTAA AGTAGGTCTG TTCGGTGGTG CCGGTGTAGG      400
15 CAAAACCGTG AACATGATGG AATTGATCAA CAACATCGCC AAAGCGCACA      450
GCGGTCTGTC CGTGTTCGCA GGTGTGGGCG AGCGTACCCG TGAAGGTAAC      500
GACTTCTACC ACGAGATGAA AGATTCCAAC GTATTGGATA AAGTGGCAAT      550
GGTTTACGGT CAGATGAACG AACCTCCGGG CAACCGTTTG CGCGTCGCAT      600
TGACCGGTTT GACCATGGCG GAATACTTCC GTGACGAAAA AGACGAAAAAC      650
20 GGTAAGGTTT GCGACGTATT GTTCTTCGTT GACAACATCT ACCGTTACAC      700
TCTGGCCGGT ACCGAAGTAT CTGCACTGTT GGGCCGTATG CCTTCTGCAG      750
TGGGTTACCA ACCGACATTG GCTGAAGAAA TGGGTCGTTT GCAAGAGCGT      800
ATTACCTCTA CCCAAACCGG TTCCATTACT TC                          832

```

25

## 2) INFORMATION FOR SEQ ID NO: 947

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 840 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Neisseria sicca*  
 (B) STRAIN: ATCC 9913

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 947

```

TCCGCGCGAT GCCATTCCGC ATGTTTACGA CGCCCTGAAA TTGGATGCAA      50
ACGGCCTGAC TTTGGAAGTA CAACAGCTTC TGGGCGACGG CGTGGTTTCGT      100
45 ACTATTGCAA TGGGTAGTTC GGACGGTCTG AAACGCGGCA TGA CTGTAAAG      150
CAATACAGAT GCGCCGATTA CTGTGCCGGT AGGTAAAGGT ACTTTGGGAC      200
GTATTGTCTGA TGTGTTGGGT ACACCTGTTG ATGAAGCAGG TCCGATTGAT      250
ACCGACAAAC ACCGTGCTAT CCATCAGACA GCTCCGAAAT TCGATGAGTT      300
GTCTGCTACT ACCGAGCTGC TGGAAACAGG CATTAAAGTG ATTGACTTGC      350
50 TGTGTCCGTT TGCCAAAGGC GGTAAAGTAG GTCTGTTCGG TGGTGCCGGT      400
GTAGGCAAAA CCGTCAACAT GATGGAATTG ATTAACAACA TCGCCAAAGC      450
GCATAGTGGT TTGTCCGTGT TCGCCGGTGT GGGGGAACGT ACCCGTGAAG      500
GTAACGACTT CTACCACGAG ATGAAAGATT CCAACGTATT GGACAAAGTG      550
GCGATGGTTT ACGGTACAGT GAACGAACCT CCGGGTAACC GTCTGCGTGT      600
55 AGCCTTGACC GGTGTTGACGA TGGCCGAATA CTTCCGTGAT GAAAAAGACG      650
AAAGCGGCAA AGGTCGCGAC GTATTGTTCT TCGTGGACAA CATTTACCGT      700
TACACTCTGG CCGGTACAGA AGTATCCGCA TTGCTCGGTC GTATGCCCTC      750
AGCAGTAGGT TACCAACCGA CATTTGGCTGA AGAAATGGGT CGTCTGCAAG      800
AGCGTATTAC CCTCTACTCA AACAGGCTCC ATTACTTCTA                      840

```

60

## 2) INFORMATION FOR SEQ ID NO: 948

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 843 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 10 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Streptococcus mitis*  
 15 (B) STRAIN: ATCC 903
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 948

	GCAGCTGGCG	ACAAGCTACC	TGAGATCAAT	AATGCACTTG	TAGTCTATAA	50
20	AAATGACGAA	AAAAAATCAA	AAATCGTCCT	TGAAGTAGCT	CTTGAGCTTG	100
	GTGATGGAGT	GGTTCGGACC	ATCGCTATGG	AATCAACGGA	TGGGTTGACT	150
	CGTGGCATGG	AAGTGCTAGA	TACTGGCCGT	CCAATTTCTG	TGCCAGTCGG	200
	CAAGAAACA	CTTGGTTCGG	TCTTTAACGT	TTTGGGAGAT	ACCATTGACT	250
	TGGATGCTCC	TTTTGCGGAT	GATGCAGAGC	GCCAGCCAAT	CCATAAGAAA	300
25	GCTCCAACCT	TTGATGAGTT	GTCTACTTCT	TCAGAGATCT	TAGAGACAGG	350
	TATCAAGGTT	ATCGACCTGT	TAGCCCCTTA	TCTGAAAGGT	GGTAAAGTTG	400
	GACTCTTCGG	TGGTGCCGGA	GTTGGTAAGA	CCGTCCTGAT	TCAAGAATTG	450
	ATCCACAACA	TTGCCCAAGA	ACACGGTGGT	ATTTCTGTAT	TTACTGGCGT	500
	TGGGGAACGT	ACCCGTGAAG	GGAATGACCT	TTATTGGGAA	ATGAAAGAGT	550
30	CTGGTGTTAT	CGAGAAAACA	GCCATGGTCT	TCGGTCAGAT	GAATGAGCCG	600
	CCAGGAGCGC	GTATGCGGGT	TGCTTTGACT	GGTTTGACGA	TTGCAGAATA	650
	CTTCCGTGAT	GTGGAAGGTC	AAGATGTCTT	GCTCTTCATT	GACAACATCT	700
	TCCGTTTCAC	GCAGGCAGGT	TCTGAAGTTT	CTGCCCTTTT	GGGTCGGATG	750
	CCGTCAGCCG	TTGGTTACCA	ACCAACACTT	GCGACAGAAA	TGGGGCAATT	800
35	GCAAGAGCGT	ATCACATCGA	CTAAGAAGGG	TTCTGTAACC	TCT	843

## 2) INFORMATION FOR SEQ ID NO: 949

- 40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 841 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 45 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Streptococcus mitis*  
 50 (B) STRAIN: ATCC 49456
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 949

55	GCAGCAGGGG	AAAAACTTCC	TGAGATTAAC	AATGCACTTG	TCGTCTACAA	50
	AAATGACGAA	AGAAAAACAA	AAATCGTCCT	TGAAGTAGCC	TTGGAGTTGG	100
	GAGATGGTAT	GGTCCGTACT	ATCGCCATGG	AATCAACAGA	TGGTTTGACT	150
	CGTGGAATGG	AAGTATTGGA	CACAGGTCGT	CCAATCTCTG	TACCAGTAGG	200
	TAAAGAAACT	TTGGGACGTG	TCTTCAATGT	TTTGGGAGAT	ACCATTGACT	250
60	TGGAAGCTCC	TTTTACAGAA	GATGCAGAGC	GTCAGCCAAT	TCATAAAAAA	300

```

      GCTCCAACCTT TTGATGAATT GTCTACCTCT TCTGAAATCC TTGAAACAGG      350
      GATTAAGGTT ATCGACCTTC TTGCCCCCTTA CCTTAAAGGT GGTAAGGTTG      400
      GACTTTTCGG TGGTGCCGGA GTTGGTAAAA CCGTCTTAAT CCAAGAATTG      450
      ATTCACAACA TTGCCCAAGA ACACGGTGGT ATTCAGTAT TTACCGGTGT      500
5     TGGGGAACGT ACTCGTGAGG GTAATGACCT TTAGTGGGAA ATGAAAGAAT      550
      CAGGTGTTAT CGAGAAAACA GCCATGGTAT TTGGTCAGAT GAATGAGCCG      600
      CCAGGAGCAC GTATGCGTGT TGCCCTAACT GGTTTGACAA TCGCCGAATA      650
      CTTCCGTGAT GTGGAAGGCC AAGACGTGCT TCTCTTTATC GATAATATCT      700
      TCCGTTTCAC TCAGGCTGGT TCAGAAGTAT CTGCCCTTTT GGGTCGTATG      750
10    CCATCAGCCG TTGGTTACCA ACCAACACTT GCTACGGAAA TGGGTCAATT      800
      GCAAGAGCGT ATTACATCAA CTAAAAAGGG TTCTGTAACC T      841

```

15 2) INFORMATION FOR SEQ ID NO: 950

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 827 bases  
 (B) TYPE: Nucleic acid  
 20 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

25 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
 (B) STRAIN: LSPQ 2583

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 950

```

30    GCTACCTGAG ATCAATAATG CACTTGTAGT CTATAAAAAT GACGAAAATA      50
      AATCAAAAAT CGTCCTTGAA GTAGCTCTTG AGCTTGGTGA TGGAGTGGTT      100
      CGGACCATCG CTATGGAATC AACGGATGGG TTGACTCGTG GCATGGAAGT      150
      GCTAGATACT GGTGCTCCAA TTTCTGTGCC AGTCGGCAA GAAACACTTG      200
35    GTCGCGTCTT TAACGTTTTG GGAGATACCA TTGACTTGGA TGCTCCTTTT      250
      GCGGATGATG CAGAGCGCCA GCCAATCCAT AAGAAAGCTC CAACCTTTGA      300
      TGAGTTGTCT ACTTCATCAG AGATCTTAGA GACAGGTATC AAGGTTATCG      350
      ACCTGTTAGC ACCTTATCTG AAAGGTGGTA AAGTCGGA CTTCGGTGGT      400
      GCCGGAGTTG GTAAGACCGT CCTGATTGAG GAATTGATCC ACAACATTGC      450
40    CCAAGAGCAT GGTGGTATTT CCGTGTTTAC CCGTGTTGGG GAACGTACCC      500
      GTGAAGGGAA TGACCTTTAC TGGGAAATGA AGGAGTCTGG CGTTATCGAG      550
      AAAACAGCCA TGGTCTTCGG TCAGATGAAT GAGCCACCAG GAGCGCGTAT      600
      GCGGGTTGCT TTGACTGGTT TGACGATTGC AGAGTACTTC CGTGATGTAG      650
      AAGGTCAAGA TGTCTTGCTC TTCATTGACA ACATCTTCCG TTTCACGCAG      700
45    GCAGGTTCTG AAGTCTCTGC CCTTTTGGGT CGGATGCCAT CAGCCGTTGG      750
      TTACCAACCA ACACTTGCGA CTGAAATGGG ACAACTCCAA GAGCGTATTA      800
      CATCGACTAA GAAAGGTTCT GTAACCT      827

```

50

2) INFORMATION FOR SEQ ID NO: 951

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 844 bases  
 55 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

60

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus oralis*  
 (B) STRAIN: ATCC 35037

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 951

```

GCAGCAGGGG AAACACTTCC TGAGATTAAT AATGCACTTG TCGTCTACAA      50
AAATGACGAA AGAAAAACAA AAATCGTCCT TGAAGTAGCC TTGGAGTTGG      100
GTGATGGTAT GGTCCGTACG ATCGCCATGG AATCAACAGA TGGTTTGACT      150
10 CGTGGAATGG AAGTTTGGGA CACAGGCCGT CCAATCTCTG TACCAGTAGG      200
TAAAGAAACT TTGGGACGTG TCTTCAACGT TTTGGGAGAT ACTATTGACT      250
TGGATGCTCC TTTCGCTGAA GACGCTGAGC GTCAGCCAAT TCATAAGAAA      300
GCTCCAACCT TTGATGAATT GTCTACCTCA TCTGAAATCT TGGAAACAGG      350
GATTAAGGTT ATCGACCTTC TTGCCCCTTA CCTTAAAGGT GGGAAAGTTG      400
15 GACTCTTCGG TGGTGCCGGA GTTGGTAAAA CTGTCTTGAT CCAAGAGTTG      450
ATTCAACAAC TTGCCCAAGA ACATGGTGGT ATTTCAGTAT TTACCGGTGT      500
TGGAGAACGT ACCCGTGAGG GGAACGACCT TTAAGTGGGAA ATGAAAGAAT      550
CAGGCGTTAT CGAGAAAACA GCCATGGTAT TTGGTCAGAT GAATGAGCCA      600
CCTGGAGCAC GTATGCGTGT TGCTCTTACT GGTTTGACAA TCGCCGAATA      650
20 CTTCCGTGAT GTAGAAGGCC AAGATGTGCT TCTCTTTATC GACAATATCT      700
TCCGTTTCAC TCAAGCTGGT TCAGAAGTAT CTGCCCTTTT GGGTCGGATG      750
CCTTCAGCCG TTGGTTACCA ACCAACACTT GCTACGGAAA TGGGTCAATT      800
GCAAGAACGT ATCACATCAA CTAAGAAGGG TTCTGTAACC TCTA          844

```

25

## 2) INFORMATION FOR SEQ ID NO: 952

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 830 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-06

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 952

```

GCAGCAGGGG AAAA ACTTCC TGAGATTAAC AATGCACTTG TCGTCTACAA      50
AAATGACGAA AGAAAAACAA AAATCGTCCT TGAAGTAGCC TTGGAGTTAG      100
45 GAGATGGTAT GGTTCGTACT ATCGCCATGG AATCAACAGA TGGGTTGACT      150
CGTGGAATGG AAGTATTGGA CACAGGTCGT CCAATCTCTG TACCAGTAGG      200
TAAAGAAACT TTGGGACGTG TCTTCAACGT TTTGGGAGAT ACCATTGATT      250
TGAAGCTCC TTTTACAGAA GACGCAGAGC GTCAGCCAAT TCATAAAAAA      300
GCTCCAACCT TTGATGAGTT GTCTACCTCT TCTGAAATCC TTGAAACAGG      350
50 GATCAAGGTT ATTGACCTTC TTGCCCCTTA CCTTAAAGGT GGTAAAGTTG      400
GACTTTTCGG TGGTGCCGGA GTTGGTAAAA CCGTCTTAAT CCAAGAATTG      450
ATTCAACAAC TTGCCCAAGA GCACGGTGGT ATTTCAGTAT TTAAGTGGTGT      500
TGGGGAACGT ACTCGTGAGG GGAATGACCT TTAAGTGGGAA ATGAAAGAAT      550
CAGGCGTTAT CGAGAAAACA GCCATGGTCT TTGGTCAGAT GAATGAGCCA      600
55 CCAGGAGCAC GTATGCGTGT TGCCCTTACT GGTTTGACAA TCGCTGAATA      650
CTTCCGTGAT GTGGAAGGCC AAGAGCTGCT TCTCTTTATC GATAATATCT      700
TCCGTTTCAC TCAGGCTGGT TCAGAAGTAT CTGCCCTTTT GGGTCGTATG      750
CCATCAGCCG TTGGTTACCA ACCAACACTT GCTACGGAAA TGGGTCAATT      800
GCAAGAACGT ATCACATCAA CCAAGAAGGG          830

```

60

## 2) INFORMATION FOR SEQ ID NO: 953

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 823 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 10 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Streptococcus pneumoniae*  
 15 (B) STRAIN: StrR-11
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 953

	AAAAACTTCC	TGAGATTAAC	AATGCACTTG	TCGTCTACAA	AAATGACGAA	50
20	AGAAAAACAA	AAATCGTCCT	TGAAGTAGCC	TTGGAGTTAG	GAGATGGTAT	100
	GGTTCGTACT	ATCGCCATGG	AATCAACAGA	TGGGTTGACT	CGTGGAATGG	150
	AAGTATTGGA	CACAGGTCGT	CCAATCTCTG	TACCAGTAGG	TAAAGAAACT	200
	TTGGGACGTG	TCTTCAACGT	TTTGGGAGAT	ACCATTGATT	TGGAAGCTCC	250
	TTTTACAGAA	GACGCAGAGC	GTCAGCCAAT	TCATAAAAAA	GCTCCAACCT	300
25	TTGATGAGTT	GTCTACCTCT	TCTGAAATCC	TTGAAACAGG	GATCAAGGTT	350
	ATTGACCTTC	TTGCCCCTTA	CCTTAAAGGT	GGTAAAGTTG	GACTTTTCGG	400
	TGGTGCCGGA	GTTGGTAAAA	CCGTCTTAAT	CCAAGAATTG	ATTCACAACA	450
	TTGCCCAAGA	GCACGGTGGT	ATTTCAGTAT	TTACTGGTGT	TGGGGAACGT	500
	ACTCGTGAGG	GGAATGACCT	TTACTGGGAA	ATGAAAGAAT	CAGGCGTTAT	550
30	CGAGAAAACA	GCCATGGTCT	TTGGTCAGAT	GAATGAGCCA	CCAGGAGCAC	600
	GTATGCGTGT	TGCCCTTACT	GGTTTGACAA	TCGCTGAATA	CTTCCGTGAT	650
	GTGGAAGGCC	AAGACGTGCT	TCTCTTTATC	GATAATATCT	TCCGTTTCAC	700
	TCAGGCTGGT	TCAGAAGTAT	CTGCCCTTTT	GGGTCGTATG	CCATCAGCCG	750
	TTGGTTACCA	ACCAACACTT	GCTACGGAAT	TGGGTCAATT	GCAAGAACGT	800
35	ATCACATCAA	CTAAGAAGGG	TTC			823

## 2) INFORMATION FOR SEQ ID NO: 954

- 40 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 844 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 45 (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:  
 50 (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 954

55	GCAGCAGGGG	AAAAACTTCC	TGAGATTAAC	AATGCACTTG	TCGTCTACAA	50
	AAATGACGAA	AGAAAAACAA	AAATCGTCCT	TGAAGTAGCC	TTGGAGTTAG	100
	GAGATGGTAT	GGTTCGTACT	ATCGCCATGG	AATCAACAGA	TGGGTTGACT	150
	CGTGGAATGG	AAGTATTGGA	CACAGGTCGT	CCAATCTCTG	TACCAGTAGG	200
	TAAAGAAACT	TTGGGACGTG	TCTTCAACGT	TTTGGGAGAT	ACCATTGACT	250
60	TGGAAGCTCC	TTTTACAGAA	GACGCAGAGC	GTCAGCCAAT	TCATAAAAAA	300

```

GCTCCAACCTT TTGATGAGTT GTCTACCTCT TCTGAAATCC TTGAAACAGG 350
GATCAAGGTT ATTGACCTTC TTGCCCCCTTA CCTTAAAGGT GGTAAAGTTG 400
GACTTTTCGG TGGTGCCGGA GTTGGTAAAA CTGTCTTAAT CCAAGAATTG 450
ATTACAACA TTGCCCAAGA GCACGGTGGT ATTTCAGTAT TTGCTGGTGT 500
5 TGGGGAACGT ACTCGTGAGG GGAATGACCT TTAAGGGAA ATGAAAGAAT 550
CAGGCGTTAT CGAGAAAACA GCCATGGTCT TTGGTCAGAT GAATGAGCCA 600
CCAGGAGCAC GTATGCGTGT TGCCCTTACT GGTTTGACAA TCGCTGAATA 650
CTTCCGTGAT GTGGAAGGCC AAGACGTGCT TCTCTTTATC GATAATATCT 700
TCCGTTTCAC TCAGGCTGGT TCAGAAGTAT CTGCCCTTTT GGGTCGTATG 750
10 CCATCAGCCG TTGGTTACCA ACCAACACTT GCTACGAAA TGGGTCAATT 800
GCAAGAACGT ATCACATCAA CCAAGAAGGG TTCTGTAACC TCTA 844

```

## 15 2) INFORMATION FOR SEQ ID NO: 955

## (i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 834 bases
(B) TYPE: Nucleic acid
20 (C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 25 (vi) ORIGINAL SOURCE:

```

(A) ORGANISM: Streptococcus pneumoniae
(B) STRAIN: StrR-05

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 955

```

30 AAAAACTTCC TGAGATTAAC AATGCACTTG TCGTCTACAA AAATGACGAA 50
AGAAAAACAA AAATCGTCCT TGAAGTAGCC TTGGAGTTAG GAGATGGTAT 100
GGTTCGTACT ATCGCCATGG AATCAACAGA TGGGTTGACT CGTGGAATGG 150
AAGTATTGGA CACAGGTCGT CCAATCTCTG TACCAGTAGG TAAAGAACT 200
35 TTGGGACGTG TCTTCAACGT TTTGGGAGAT ACCATTGATT TGGAAGCTCC 250
TTTTACAGAA GACGCAGAGC GTCAGCCAAT TCATAAAAAA GCTCCAACCT 300
TTGATGAGTT GTCTACCTCT TCTGAAATCC TTGAAACAGG GATCAAGGTT 350
ATTGACCTTC TTGCCCCCTTA CCTTAAAGGT GGTAAAGTTG GACTTTTCGG 400
TGGTGCCGGA GTTGGTAAAA CCGTCTTAAT CCAAGAATTG ATTCACAACA 450
40 TTGCCCAAGA GCACGGTGGT ATTTCAGTAT TTAAGGGTGT TGGGGAACGT 500
ACTCGTGAGG GGAATGACCT TTAAGGGAA ATGAAAGAAT CAGGCGTTAT 550
CGAGAAAACA GCCATGGTCT TTGGTCAGAT GAATGAGCCA CCAGGAGCAC 600
GTATGCGTGT TGCCCTTACT GGTTTGACAA TCGCTGAATA CTTCCGTGAT 650
GTGGAAGGCC AAGACGTGCT TCTCTTTATC GATAATATCT TCCGTTTCAC 700
45 TCAGGCTGGT TCAGAAGTAT CTGCCCTTTT GGGTCGTATG CCATCAGCCG 750
TTGGTTACCA ACCAACACTT GCTACGAAA TGGGTCAATT GCAAGAACGT 800
ATCACATCAA CCAAGAAGGG TTCTGTAACC TCTA 834

```

## 50 2) INFORMATION FOR SEQ ID NO: 956

## (i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 495 bases
55 (B) TYPE: Nucleic acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

60

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Babesia microti*  
 (B) STRAIN: Persing-1

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 956

```

TTGTATATCA CAGGCACTCA GCAAATATTC CGATACTGAC GTAATTATAT 50
ACGTGGGTTG TGGTGAACGT GGGAAATGAAA TGGCTGAGAT TCTATGCGAA 100
TTCCCTGAAC TATCTACTGT AGTTAATGAT GAAAAGGTGG CCATTATGGA 150
10 ACGTACATGC TTAGTTGCCA ATACTTCTAA TATGCCAGTG GCCGCTAGAG 200
AAGCTAGTAT ATACACTGGT ATTACAATTG CTGAATATTT CCGTGATATG 250
GGTTACAAC TACTCTTAT GGCCGATTCC ACTAGCCGAT GGGCAGAGGC 300
TCTAAGGGAA ATTTCTGGTA GATTGGCTGA AATGCCTGCA GATTCTGGCT 350
ATCCGGCCTA TTTATCGTCA AGGTTGTCAG CTTTTTATGA ACGTGCAGGT 400
15 GGGATAACTG TCTAATTAAT TTAGGCTTGA TTAAGTGCTT AGGTTACCA 450
ACACGAACCG GATCTATTAC GGTGTAGGA GCAGTTTCTC CACCA 495

```

## 20 2) INFORMATION FOR SEQ ID NO: 957

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 469 bases  
 (B) TYPE: Nucleic acid  
 25 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Entamoeba histolytica*  
 (B) STRAIN: HM1-1MSS

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 957

```

35 AGTTATTTCA CAAGCATTA GATAATATAG TAATTCAGAT GTTATTATTT 50
ATGTAGGATG TGGTGAACGA GGAAATGAAA TGGCAGAAGT TCTTCGAGAT 100
TTTCCAGCTC TTTCTATTAA AGTAGGAGAT AAAGAAGAAT CTATTATGAC 150
AAGAACAGCA CTTGTTGCTA ATACATCTAA TATGCCTGTT GCAGCACGTG 200
40 AAGCATCAAT TTATACTGGA ATTACATTAT CAGAATATTA TAGAGATATG 250
GGATATAATG TTGCTATGAT GGCAGATTCA ACATCAAGAT GGGCTGAAGC 300
ACTTAGAGAA ATTTCAGGAC GTCTTGCAGA AATGCCAGCT GATTCTGGAT 350
ATCCAGCATA TCTTGCAGCA CGTTTAGCAT CATTTTATGA ACGTGCAGGT 400
ATGGTTGAAT GTTTAGGATC ACCAAAAAGA ATAGGGTCAG TTTCTATTGT 450
45 AGGAGCTGTT TCACCACCT 469

```

## 2) INFORMATION FOR SEQ ID NO: 958

50

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 452 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 55 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- 60 (A) ORGANISM: *Fusobacterium nucleatum* subsp. *polymorphum*

(B) STRAIN: ATCC 10953

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 958

```

5  TACAACACCA ACTTGCTAAA TGGGCAGATG CAGAAGTAGT TGTTTATGTT      50
   GGTGTGTTGGG AACGTGGAAA TGAAATGACC GATGTACTTA TGGAATTCCC      100
   AGAAATTATT GACCCTAAGA CAGGACAATC TTTAATGAAG AGAACAGTTC      150
   TTATAGCTAA TACTTCTAAT ATGCCAGTTG CTGCTCGTGA GGCTTCAATC      200
   TATACTGGTA TAACTATTGC AGAATATTTT AGAGATATGG GATATTCAGT      250
10  GGCACCTTATG GCAGATTCAA CAAGTCGTTG GGCAGAAGCA CTTCGTGAAA      300
   TGTCAGGACG TTTGGAAGAA ATGCCAGGTG ATGAAGGATA TCCAGCATAT      350
   CTATCAAGTA GAATAGCAGA GTTTTATGAA AGAGCAGGGC TTGTTGAATG      400
   TCTAGGTAAT GGAGAAGAAG GAGCATTAAC TGTAATTGGA GCAGTATCTC      450
   CA
15

```

2) INFORMATION FOR SEQ ID NO: 959

```

20  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 469 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear
25
   (ii) MOLECULE TYPE: Genomic DNA
   (vi) ORIGINAL SOURCE:
30  (A) ORGANISM: Leishmania aethiopica
      (B) STRAIN: ATCC 50119

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 959

```

35  TGTAATTAGT CAGGCCCTCT CCAAGTACTC CAACTCCGAC TCGTCATCT      50
   ATGTCGGCTG CGGCGAGCGC GGTAATGAGA TGGCCGAGGT GTCATGGAG      100
   TTCCCGACCC TGACGACCGT GATCAATGGT CGCGAGGAGT CGATCATGAA      150
   GCGCACCTGC CTCGTGGCGA ACACTTCGAA CATGCCAGTC GCAGCCCGTG      200
   AGGCCTCTAT TTACACCGGC ATCACCTGG CCGAGTACTA CCGTGATATG      250
   GGCAAGCATA TCGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCCGAGGC      300
40  GCTTCGTGAG ATTTCGGGTC GTCTGGCGGA GATGCCGGCC GATGGTGGCT      350
   ACCCTGCCTA TCTCAGCGCT CGTCTCGCCT CCTTCTACGA GCGCGCCGGC      400
   CTCGTACCT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATTGT      450
   CGGTGCCGTG TCTCCGCCG
45

```

2) INFORMATION FOR SEQ ID NO: 960

```

50  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 469 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear
55
   (ii) MOLECULE TYPE: Genomic DNA
   (vi) ORIGINAL SOURCE:
60  (A) ORGANISM: Leishmania tropica
      (B) STRAIN: ATCC 30815

```



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 960

```

      TGTAATTAGT CAGGCCCTCT CCAAGTACTC CAACTCCGAC TCGGTCATCT      50
      ATGTCGGCTG CGGCGAGCGC GGTAATGAGA TGGCCGAGGT GCTCATGGAG      100
5     TTCCCGACCC TGACGACCGT GATCGATGGT CGCGAGGAGT CGATCATGAA      150
      GCGCACCTGC CTCGTGGCGA ACACCTCGAA CATGCCAGTC GCAGCCCGTG      200
      AGGCCTCTAT TTACACCGGC ATCACCCTGG CCGAGTACTA CCGTGATATG      250
      GGCAAGCATA TCGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCCGAGGC      300
      GCTTCGTGAG ATTTTCGGGTC GTTTGGCGGA GATGCCGGCC GATGGTGGCT      350
10    ACCCTGCCTA TCTCAGCGCT CGTCTCGCCT CCTTCTACGA GCGCGCCGGC      400
      CTCGTACCTT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATTGT      450
      CGGTGCCGTG TCTCCGCCG                                469

```

15

## 2) INFORMATION FOR SEQ ID NO: 961

## (i) SEQUENCE CHARACTERISTICS:

```

20    (A) LENGTH: 469 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

25

## (vi) ORIGINAL SOURCE:

```

      (A) ORGANISM: Leishmania guyanensis
      (B) STRAIN: ATCC 50126

```

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 961

```

      TGTTCATCAGT CAGGCCCTCT CCAAGTACTC CAACTCCGAC TGTGTCATCT      50
      ATGTCGGCTG CGGTGAACGC GGTAACGAGA TGGCCGAGGT GCTCATGGAG      100
      TTCCCGACCC TGACGACTGT GATCGATGGT CGCGAAGAGT CCATCATGAA      150
35    GCGCACCTGC CTCGTGGCGA ACACCTCGAA CATGCCCGTC GCAGCCCGTG      200
      AGGCCTCTAT TTATACCGGC ATCACCCTTG CTGAGTACTA CCGTGATATG      250
      GGCAAGCACA TTGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCCGAGGC      300
      GCTGCGTGAG ATTTTCGGGTC GATTGGCGGA GATGCCGGCT GATGGTGGCT      350
      ACCCTGCCTA CCTCAGCGCC CGCCTCGCCT CCTTCTACGA GCGCGCCGGT      400
40    CTCGTACCTT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATCGT      450
      CGGTGACAGT TCTCCACCG                                469

```

## 45 2) INFORMATION FOR SEQ ID NO: 962

## (i) SEQUENCE CHARACTERISTICS:

```

50    (A) LENGTH: 469 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 55 (vi) ORIGINAL SOURCE:

```

      (A) ORGANISM: Leishmania donovani
      (B) STRAIN: ATCC 50212

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 962

60

```

TGTCATTAGT CAGGCCCTCT CCAAGTACTC CAACTCCGAT TGCATCATCT      50
ATGTCGGCTG CGGCGAGCGC GGTAATGAGA TGGCCGAGGT GCTCATGGAG      100
TTCCCGACCC TGACGACCGT GATCGATGGC CGCGAGGAGT CGATCATGAA      150
GCGCACCTGC CTCGTGGCGA ACACCTCGAA CATGCCAGTC GCAGCCCGTG      200
5 AGGCCTCTAT TTACACCGGC ATCACCTTGG CCGAGTACTA CCGTGATATG      250
GGCAAGCATA TCGCCATGAT GGCTGACTCG ACGTCTCGCT GGGCCGAGGC      300
GCTTCGTGAG ATTTCCGGTC GTCTGGCGGA GATGCCGGCC GATGGTGGCT      350
ACCCCGCCTA CCTCAGCGCT CGTCTCGCCT CCTTCTACGA GCGCGCCGGC      400
CTCGTCACCT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATCGT      450
10 CGGTGCCGTG TCTCCACCG                                     469

```

## 2) INFORMATION FOR SEQ ID NO: 963

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 469 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
20 (D) TOPOLOGY: Linear

    (ii) MOLECULE TYPE: Genomic DNA

    (vi) ORIGINAL SOURCE:
25 (A) ORGANISM: Leishmania hertigi
    (B) STRAIN: ATCC 50125

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 963

```

30 TGTGATCAGT CAGGCCCTTT CCAAGTACTC CAACTCGGAC TGCATCATCT      50
ACGTCGGCTG TGGTGAGCGC GGGAACGAGA TGGCCGAGGT GCTCATGGAT      100
TTCCCGACTT TGACGACTGT GATCGATGGT CGCGAGGAGT CCATCATGAA      150
GCGCACCTGC CTCGTGGCGA ACACCTCCAA CATGCCAGTT GCAGCCCGTG      200
AGGCTTCTAT CTATACCGGC ATCACACTGG CTGAGTACTA TCGTGATATG      250
35 GGCAAGCACA TTGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCCGAGGC      300
GTTGCGTGAG ATTTCCGGTC GGCTGGCGGA GATGCCGGCC GATGGTGGTT      350
ACCCCGCCTA CCTCAGTGCC CGTCTCGCCT CCTTCTACGA GCGCGCTGGC      400
CTCGTGACCT GTATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACAATTGT      450
40 TGGTGCGGTG TCTCCACCG                                     469

```

## 2) INFORMATION FOR SEQ ID NO: 964

```

45 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 469 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
50 (D) TOPOLOGY: Linear

    (ii) MOLECULE TYPE: Genomic DNA

    (vi) ORIGINAL SOURCE:
55 (A) ORGANISM: Leishmania mexicana
    (B) STRAIN: ATCC 50156

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 964

```

60 GGTCATTAGT CAGGCCCTTT CCAAGTACTC CAACTCTGAC TGCATCATCT      50
ACGTCGGCTG CGGCGAGCGC GGTAATGAGA TGGCCGAGGT GCTCATGGAG      100

```

TTCCCGACCC TGACGACCAT GATCGATGGT CGGGAGGAGT CGATCATGAA 150  
 GCGCACCTGC CTCGTGGCGA ACACCTCGAA CATGCCCGTC GCAGCCCGTG 200  
 AGGCCTCTAT CTACACCGGC ATCACCCCTCG CCGAGTACTA CCGTGATATG 250  
 GGCAAGCATA TCGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCTGAGGC 300  
 5 GCTTCGTGAG ATTTTCGGGTC GTCTGGCGGA GATGCCGGCC GATGGTGGCT 350  
 ACCCCGCCTA CCTCAGCGCT CGTCTCGCCT CCTTCTACGA GCGCGCCGGC 400  
 CTCGTACCT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATCGT 450  
 CCGTGCCGTG TCTCCGCCG 469

10

## 2) INFORMATION FOR SEQ ID NO: 965

(i) SEQUENCE CHARACTERISTICS:  
 15 (A) LENGTH: 469 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
 (ii) MOLECULE TYPE: Genomic DNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Leishmania tropica*  
 (B) STRAIN: ATCC 50129

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 965

TGTAATTAGT CAGGCCCTCT CCAAGTACTC CAACTCCGAC TCGGTCATCT 50  
 ATGTCGGCTG CGGCGAGCGC GGTAAATGAGA TGGCCGAGGT GCTCATGGAG 100  
 30 TTCCCGACCC TGACGACCGT GATCGATGGT CGCGAGGAGT CGATCATGAA 150  
 GCGCACCTGC CTCGTGGCGA ACACCTTCGAA CATGCCAGTC GCAGCCCGTG 200  
 AGGCCTCTAT TTACACCGGC ATCACCCCTGG CCGAGTACTA CCGTGATATG 250  
 GGCAAGCATA TCGCCATGAT GGCCGACTCG ACGTCTCGCT GGGCCGAGGC 300  
 GCTTCGTGAG ATTTTCGGGTC GTTTGGCGGA GATGCCGGCC GATGGTGGCT 350  
 35 ACCCTGCCTA TCTCAGCGCT CGTCTCGCCT CCTTCTACGA GCGCGCCGGC 400  
 CTCGTACCT GCATCGGCGG GCCGAAGCGC CAGGGCTCCG TCACGATTGT 450  
 CCGTGCCGTG TCTCCGCCG 469

40

## 2) INFORMATION FOR SEQ ID NO: 966

(i) SEQUENCE CHARACTERISTICS:  
 45 (A) LENGTH: 449 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
 (ii) MOLECULE TYPE: Genomic DNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Peptostreptococcus anaerobius*  
 (B) STRAIN: ATCC 27337

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 966

CACCAGTTCG CCAAGTGGGC AGATGCTCAG ATAGTTGTAT ACGTTGGTTG 50  
 TGGAGAACGT GGTAACGAGA TGACAGACGT TCTAAATGAA TTCCAGAAC 100  
 TGATTGACCC TCATACAGGC GAATCTCTAA TGAAGAGAAC AGTTCTTATA 150  
 60 GCTAATACGT CAAATATGCC AGTTGCAGCC AGAGAGGCAA GTATATATAC 200

	AGGTATTACA	ATAGCTGAAT	ATTTTAGAGA	CATGGGATAT	TCAGTAGCGG	250
	TAATGGCCGA	CTCTACATCA	AGATGGGCAG	AGGCCCTAAG	AGAGATGTCA	300
	GGTCGTCTAG	AAGAAATGCC	TGGTGATGAA	GGTTATCCAG	CCTATCTAGG	350
	TTCTAGAGCT	GCAGAGTTCT	ATGAAAGAGC	AGGTAAGGTA	ATATGTAAGG	400
5	GTAGCGATAA	TAGAGAGGGA	GCCCTTACAA	TAATAGGTGC	CGTGTCCACC	449

## 2) INFORMATION FOR SEQ ID NO: 967

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 826 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bordetella pertussis*
- (B) STRAIN: ATCC 9797

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 967

25	CTGGTGGTGT	CGGCCGCAGA	CGGCCCGATG	CCGCAGACGC	GCGAGCACAT	50
	TTTGCTGTCT	CGCCAGGTTG	GCGTGCCGTA	CATCATCGTG	TTCCTGAACA	100
	AGGCGGACAT	GGTTGATGAC	GCGGAGCTGC	TCGAGCTGGT	GAAGATGGAA	150
	GTCCGCGAAC	TGCTGAGCAA	GTACGATTTT	CCGGGCGATG	ACACGCCGAT	200
	CGTGAAGGGT	TCGGCCAAGC	TGGCGCTGGA	AGGCGACAAG	GGCGAACTGG	250
30	GCGAGCAGGC	GATTCTGTCT	CTGGCGCAAG	CGCTGGACAC	GTACATTCCG	300
	ACGCCGGAGC	GCGCGGTCGA	CGGTGCGTTC	CTGATGCCCG	TGGAAGACGT	350
	GTTCTCGATC	TCGGGCCCGT	GCACGGTGGT	GACTGGCCGT	ATCGAGCGCG	400
	GCGTGGTGAA	GGTTGGCGAG	GAAATCGAAA	TCGTGGGCAT	CAAGCCGACG	450
	GTGAAGACGA	CCTGCACGGG	CGTGGAGATG	TTCCGCAAGC	TGCTGGACCA	500
35	GGGCCAGGCG	GGCGACAACG	TGGGTATCTT	GCTGCGCGGC	ACCAAGCGTG	550
	AAGACGTCGA	GCGTGGCCAG	GTGCTGGCCA	AGCCGGGTTC	GATCAACCCG	600
	CACACGGACT	TCACGGCCGA	GGTGATACAT	CTGTCCAAGG	AAGAGGGTGG	650
	CCGTCACACG	CCGTTCTTCA	ACGGCTATCG	TCCGCAGTTC	TACTTCCGCA	700
	CGACGGACGT	GACCGGCACG	ATCGACCTGC	CGCGGGACAA	GGAAATGGTG	750
40	CTGCCGGGCG	ACAACGTGTC	GATGACCGTC	AAGCTGCTGG	CCCCGATCGC	800
	CATGGAAGAA	GGTCTGCGTT	TCGCCA			826

## 2) INFORMATION FOR SEQ ID NO: 968

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 817 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bordetella pertussis*
- (B) STRAIN: BD180

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 968

5 CGATCCTGGT GGTGTCGGCC GCAGACGGCC CGATGCCGCA GACGCGCGAG 50  
 CACATTTTGC TGTCGCGCCA GGTGCGCGTG CCGTACATCA TCGTGTTTCCT 100  
 GAACAAGGCG GACATGGTTG ATGACGCGGA GCTGCTCGAG CTGGTGGAGA 150  
 TGGAAGTCCG CGAACTGCTG AGCAAGTACG ATTTCCCGGG CGATGACACG 200  
 CCGATCGTGA AGGGTTCGGC CAAGCTGGCG CTGGAAGGCG ACAAGGGCGA 250  
 ACTGGGCGAG CAGGCGATTG TGTCGCTGGC GCAAGCGCTG GACACGTACA 300  
 TTCCGACGCC GGAGCGCGCG GTCGACGGTG CGTTCCTGAT GCCGGTGGAA 350  
 GACGTGTTCT CGATCTCGGG CCGTGGCACG GTGGTGAAGT GCCGTATCGA 400  
 GCGCGGCGTG GTGAAGGTTG GCGAGGAAAT CGAAATCGTG GGCATCAAGC 450  
 10 CGACGGTGAA GACGACCTGC ACGGGCGTGG AGATGTTCCG CAAGCTGCTG 500  
 GACCAAGGCC AGGCGGGCGA CAACGTGGGT ATCTTGCTGC GCGGCACCAA 550  
 GCGTGAAAGC GTCGAGCGTG GCCAGGTGCT GGCCAAGCCG GGTTCGATCA 600  
 ACCCGCACAC GGACTTCACG GCCGAGGTGT ACATTCTGTC CAAGGAAGAG 650  
 GGTGGCCGTC ACACGCCGTT CTTCAACGGC TATCGTCCGC AGTTCTACTT 700  
 15 CCGCACGACG GACGTGACCG GCACGATCGA CCTGCCGGCG GACAAGGAAA 750  
 TGGTGCTGCC GGGCGACAAC GTGTCGATGA CCGTCAAGCT GCTGGCCCCG 800  
 ATCGCCATGG AAGAAGG 817

20

## 2) INFORMATION FOR SEQ ID NO: 969

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 637 bases  
 25 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
  
 (ii) MOLECULE TYPE: Genomic DNA  
 30 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Enterococcus columbae*  
 (B) STRAIN: ATCC 51263

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 969

CCTATGCCAC AAACCTCGTGA ACACATTCTT TTATCACGTA ACGTTGGTGT 50  
 GCCATACATC GTTGTTTTCT TAAACAAAGT TGATATGGTT GACGACGAAG 100  
 AATTATTAGA ATTAGTTGAA ATGGAAGTTC GTGACTTATT AACTGAATAT 150  
 40 GACTTCCCAG GAGACGATGT TCCTGTAATC GCTGGTTCTG CATTAAAAGC 200  
 TTTAGAAGGC GACCCTGCTT ACGAAGAAAA AATCTTAGAA TTAATGGCTG 250  
 CAGTTGACGA ATACATCCCA ACTCCAGAAC GTGACAACGA CAAACCATTC 300  
 ATGATGCCAG TTGAAGACGT GTTCTCAATT ACTGGTCGTG GACTGTGTC 350  
 TACAGGTCGT GTTGAACGTG GACAAGTTCG TGTGGTGAC GAAGTTGAAA 400  
 45 TCGTTGGTAT CGCTGACGAA ACTTCTAAAA CAACAGTTAC TGGTGTGAA 450  
 ATGTTCCGTA AATTATTAGA TTACGCTGAA GCTGGAGACA ACATCGGTGC 500  
 ATTATTACGT GGTGTGGCTC GTGAAGACAT CCAACGTGGT CAAGTATTAG 550  
 CTAAACCAGG TTCAATCACT CCACATACAA AATTCACTGC TGAAGGTAC 600  
 50 GTTTTAACTA AAGAAGAAGG TGGACGTCAT ACTCCAT 637

## 2) INFORMATION FOR SEQ ID NO: 970

- 55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 634 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 60 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus flavescens*

(B) STRAIN: ATCC 49997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 970

10	TATGCCTCAA	ACACGTGAAC	ACATCTTGTT	ATCACGTAAC	GTTGGTGTAC	50
	CATACATCGT	TGTTTTCTTA	AACAAAATGG	ATATGGTTGA	TGACGAAGAA	100
	TTACTAGAAT	TAGTTGAAAT	GGAAGTTCGT	GACTTATTGT	CAGAATATGA	150
	CTTCCCAGGC	GACGATGTTT	CTGTAATCGC	TGGTTCTGCT	TTGAAAGCTC	200
	TTGAAGGCGA	TGCTTCATAC	GAAGAAAAAA	TCATGGAATT	AATGGCTGCA	250
	GTTGACGAAT	ACGTTCCAAC	TCCAGAACGT	GACACTGACA	AACCATTCAT	300
15	GATGCCAGTC	GAAGACGTAT	TCTCAATCAC	TGGACGTGGT	ACTGTTGCTA	350
	CAGGCCGTGT	TGAACGTGGA	CAAGTTCGCG	TTGGTGACGA	AGTTGAAATC	400
	GTTGGTATTG	CTGAAGAAAC	TGCTAAAACA	ACTGTAACGT	GTGTTGAAAT	450
	GTTCCGTAAA	TTGTTAGACT	ATGCTGAAGC	AGGGGATAAC	ATTGGTGCAAT	500
	TGCTACGTGG	GGTTGCTCGT	GAAGACATCC	AACGTGGACA	AGTATTAGCT	550
20	AAAGCTGGTA	CAATCACACC	TCATACAAAA	TTTAAAGCTG	AAGTTTACGT	600
	TTTAACAAAA	GAAGAAGGTG	GACGTCACAC	TCCA		634

25 2) INFORMATION FOR SEQ ID NO: 971

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 787 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: StrR-55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 971

40	GGACCAATGC	CACAACTCG	TGAGCACATC	CTTCTTTCAC	GTCAGGTG	50
	TGTTAAACAC	CTTATCGTCT	TCATGAACAA	AGTTGACTTG	GTTGACGACG	100
	AAGAATTGCT	TGAATTGGTT	GAAATGGAAA	TCCGTGACCT	ATTGTCAGAA	150
	TACGACTTCC	CAGGTGACGA	TCTTCCAGTT	ATCCAAGGTT	CAGCACTTAA	200
45	AGCTCTTGAA	GGTGACTCTA	AATACGAAGA	CATCGTTATG	GAATTGATGA	250
	ACACAGTTGA	TGAGTATATC	CCAGAACCAG	AACGTGACAC	TGACAAACCA	300
	TTGCTTCTTC	CAGTCGAGGA	CGTATFTCTA	ATCACTGGAC	GTGGTACAGT	350
	TGCTTCAGGA	CGTATCGACC	GTGGTATCGT	TAAAGTCAAC	GACGAAATCG	400
	AAATCGTTGG	TATCAAAGAA	GAAACTCAA	AAGCAGTTGT	TACTGGTGTT	450
50	GAAATGTTCC	GTAACAACCT	TGACGAAGGT	CTTGCTGGAG	ATAACGTAGG	500
	TGTCCTTCTT	CGTGGTGTTT	AACGTGATGA	AATCGAACGT	GGACAAGTTA	550
	TCGCTAAACC	AGGTTCAATC	AACCCACACA	CTAAATTCAA	AGGTGAAGTC	600
	TACATCCTTA	CTAAAGAAGA	AGGTGGACGT	CACACTCCAT	TCTTCAACAA	650
	CTACCGTCCA	CAATTCTACT	TCCGTACTAC	TGACGTTACA	GGTTCAATCG	700
55	AACTTCCAGC	AGGTACTGAA	ATGGTAATGC	CTGGTGATAA	CGTGACAATC	750
	GACGTTGAGT	TGATTACACC	AATCGCCGTA	GAACAAG		787

60 2) INFORMATION FOR SEQ ID NO: 972

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 803 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 35401

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 972

```

15 TAGTTGCTGC GACTGACGGC CCGATGCCGC AGACTCGTGA GCACATCCTG      50
   CTGGGTCGTC AGGTAGGCGT TCCGTACATC ATCGTGTTCC TGAACAAATG      100
   CGACATGGTT GATGACGAAG AGCTGCTGGA ACTGGTTGAA ATGGAAGTTC      150
   GTGAACTTCT GTCTCAGTAC GACTTCCCGG GCGACGACAC TCCGATCGTT      200
20 CGTGGTTCTG CTCTGAAAGC GCTGGAAGGC GACGCAGAGT GGAAGCGAA      250
   AATCCTGGAA CTGGCTGGCT TCCTGGATTG TTACATTCCG GAACCAGAGC      300
   GTGCGATTGA CAAGCCGTTT CTGCTGCCGA TCGAAGACGT ATTCTCCATC      350
   TCCGGTCGTG GTACCGTTGT TACCGGTCGT GTAGAACGCG GTATCATCAA      400
   AGTTGGTGAA GAAGTTGAAA TCGTTGGTAT CAAAGAGACT CAGAAGTCTA      450
25 CCTGTACTGG CGTTGAAATG TTCCGCAAAC TGCTGGACGA AGGCCGTGCT      500
   GGTGAGAACG TAGGTGTTCT GCTGCGTGGT ATCAAACGTG AAGAAATCGA      550
   ACGTGGTCAG GTACTGGCTA AGCCGGGCAC CATCAAGCCG CACACCAAGT      600
   TCGAATCTGA AGTGATACAT CTGTCCAAAG ATGAAGGCGG CCGTCATACT      650
   CCGTTCTTCA AAGGCTACCG TCCGCAGTTC TACTTCCGTA CTACTGACGT      700
30 GACTGGTACC ATCGAACTGC CGGAAGGCGT AGAGATGGTA ATGCCGGGCG      750
   ACAACATCAA AATGGTTGTT ACCCTGATCC ACCCGATCGC GATGGACGAC      800
   GGT

```

## 2) INFORMATION FOR SEQ ID NO: 973

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: ATCC 43895

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 973

```

50 ACTGACGGCC CGATGCCGCA GACTCGTGAG CACATCCTGC TGGGTCGTCA      50
   GGTAGGCGTT CCGTACATCA TCGTGTTTCCT GAACAAATGC GACATGGTTG      100
   ATGACGAAGA GCTGCTGGAA CTGGTTGAAA TGGAAGTTTCG TGAACCTCTG      150
55 TCTCAGTACG ACTTCCCGGG CGACGACACT CCGATCGTTC GTGGTTCTGC      200
   TCTGAAAGCG CTGGAAGGCG ACGCAGAGTG GGAAGCGAAA ATCCTGGAAC      250
   TGGCTGGCTT CCTGGATTCC TACATTCCGG AACCAGAGCG TGCRAATTGAC      300
   AAGCCGTTCC TGCTGCCGAT CGAAGACGTA TTCTCCATCT CCGGTCGTGG      350
   TACCGTTGTT ACCGGTCGTG TAGAACGCGG TATCATCAAA GTTGGTGAAG      400
60 AAGTTGAAAT CGTTGGTATC AAAGAGACTC AGAAGTCTAC CTGTACTGGC      450

```

	GTTGAAATGT	TCCGCAAAC	GCTGGACGAA	GGCCGTGCTG	GTGAGAACGT	500
	AGGTGTTCTG	CTGCGTGGTA	TCAAACGTGA	AGAAATCGAA	CGTGGTCAGG	550
	TACTGGCTAA	GCCGGGCACC	ATCAAGCCGC	ACACCAAGTT	CGAATCTGAA	600
	GTGTACATTC	TGTCCAAAGA	TGAAGGCGGC	CGTCATACTC	CGTTCTTCAA	650
5	AGGCTACCGT	CCGCAGTTCT	ACTTCCGTAC	TACTGACGTG	ACTGGTACCA	700
	TTGAACTGCC	GGAAGGCGTA	GAGATGGTAA	TGCCGGGCGA	CAACATCAAA	750
	ATGTTGTTA	CC				762

10

## 2) INFORMATION FOR SEQ ID NO: 974

## (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 804 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

20

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 11775

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 974

	CCTGGTAGTT	GCTGCGACTG	ACGGCCCGAT	GCCGCAGACT	CGTGAGCACA	50
	TCCTGCTGGG	TCGTCAGGTA	GGCGTTCGGT	ACATCATCGT	GTTCCCTGAAC	100
	AAATGCGACA	TGGTTGATGA	CGAAGAGCTG	CTGGAAGTGG	TTGAAATGGA	150
30	AGTTCGTGAA	CTTCTGTCTC	AGTACGACTT	CCCGGGCGAC	GACACTCCGA	200
	TCGTTCTGG	TTCTGCTCTG	AAAGCGCTGG	AAGGCGACGC	AGAGTGGGAA	250
	GCGAAAATCC	TGGAAGTGGC	TGGCTTCCTG	GATTCCTACA	TTCCGGAACC	300
	AGAGCGTGCG	ATTGACAAGC	CGTTCCTGCT	GCCGATCGAA	GACGTATTCT	350
	CCATCTCCGG	TCGTGGTACC	GTTGTTACCG	GTCGTGTAGA	ACGCGGTATC	400
35	ATCAAAGTTG	GTGAAGAAGT	TGAAATCGTT	GGTATCAAAG	AGACTCAGAA	450
	GTCTACCTGT	ACTGGCGTTG	AAATGTTCCG	CAAAGTGGT	GACGAAGGCC	500
	GTGCTGGTGA	GAACGTAGGT	GTTCTGCTGC	GCGGTATCAA	ACGTGAAGAA	550
	ATCGAACGTG	GTCAGGTACT	GGCTAAGCCG	GGCACCATCA	AGCCGCACAC	600
	CAAGTTCGAA	TCTGAAGTGT	ACATTCTGTC	CAAAGATGAA	GGCGGTCGTC	650
40	ATACTCCGTT	CTTCAAAGGC	TACCGTCCGC	AGTTCTACTT	CCGTACTACT	700
	GACGTGACTG	GTACCATCGA	ACTGCCGGAA	GGCGTAGAGA	TGGTAATGCC	750
	GGGCGACAAC	ATCAAAATGG	TTGTTACCCT	GATCCACCCG	ATCGCGATGG	800
	ACGA					804

45

## 2) INFORMATION FOR SEQ ID NO: 975

## (i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 804 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

55

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*
- (B) STRAIN: ATCC 25922

60



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 975

```

5  GCGATCCTGG TAGTTGCTGC GACTGACGGC CCGATGCCGC AGACTCGTGA 50
   GCACATCCTG CTGGGTCGTC AGGTAGGCGT TCCGTACATC ATCGTGTTCC 100
   TGAACAAATG CGACATGGTT GATGACGAAG AGCTGCTGGA ACTGGTTGAA 150
   ATGGAAGTTC GTGAACTTCT GTCTCAGTAC GACTTCCCCG GCGACGACAC 200
   TCCGATCGTT CGTGGTTCTG CTCTGAAAGC GCTGGAAGGC GACGCAGAGT 250
   GGGAAGCGAA AATCCTGGAA CTGGCTGGCT TCCTGGATTG YTACATTCCG 300
   GAACCAGAGC GTGCGATTGA CAAGCCGTTT CTGCTGCCGA TCGAAGACGT 350
10 ATTCTCCATC TCCGGTCGTG GTACCGTTGT TACCGGTCGT GTAGAACGCG 400
   GTATCATCAA AGTTGGTGAA GAAGTTGAAA TCGTTGGTAT CAAAGAGACT 450
   CAGAAGTCTA CCTGTACTGG CGTTGAAATG TTCCGCAAAC TGCTGGACGA 500
   AGGCCGTGCT GGTGAGAACG TAGGTGTTCT GCTGCGTGGT ATCAAACGTG 550
   AAGAAATCGA ACGTGGTCAG GTACTGGCTA AGCCGGGCAC CATCAAGCCG 600
15 CACACCAAGT TCGAATCTGA AGTGATACAT CTGTCCAAAG ATGAAGGCGG 650
   TCGTCATACT CCGTTCTTCA AAGGCTACCG TCCGCAGTTC TACTTCCGTA 700
   CTACTGACGT GACTGGTACC ATCGAAGTGC CGGAAGGTGT AGAGATGGTA 750
   ATGCCGGGCG ACAACATCAA AATGGTTGTT ACCCTGATCC ACCCGATCGC 800
   GATG 804
20

```

## 2) INFORMATION FOR SEQ ID NO: 976

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 825 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 30 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Mycobacterium avium*
  - (B) STRAIN: Mavi-1
- 35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 976

```

40 GGCGCGATCC TGGTGGTTCG CGCCACCGAC GGCCCGATGC CGCAGACCCG 50
   TGAGCACGTG CTGCTCGCCC GTCAGGTCGG TGTGCCCTAC ATCCTGGTCG 100
   CCCTGAACAA GGCCGACATG GTCGACGACG AGGAGCTCCT CGAGCTCGTC 150
   GAGATGGAGG TCCGCGAGCT GCTGGCCGCC CAGGAGTTCG ACGAGGACGC 200
   CCCGGTGGTG CGGGTCTCGG CGCTCAAGGC GCTCGAGGGC GACGCCAAGT 250
   GGGTCGAGTC CGTCGAGCAG CTGATGGAGG CCGTCGACGA GTCGATCCCG 300
45 GACCCGGTCC GCGAGACGGA GAAGCCGTTT CTGATGCCGG TGGAGGACGT 350
   CTTCAACATC ACCGGGCGTG GCACCGTGGT CACCGGTCGT GTCGAGCGCG 400
   GTGTGATCAA CGTGAACGAG GAAGTCGAGA TCGTCGGCAT CCGCCCGACC 450
   AGCACCAAGA CCACGGTCAC CGGTGTGGAR ATGTTCCGCA AGCTGCTCGA 500
   CCAGGGCCAG GCCGGTGACA ACGTCGGTCT GCTGCTGCGT GGTATCAAGC 550
50 GTGAGGACGT CGAGCGCGGC CAGGTCGTGA CCAAGCCCGG CACCACCACG 600
   CCGCACACCG AGTTCGAGGG CCAGGTCTAC ATCCTGTCCA AGGACGAGGG 650
   CGGCCGGCAC ACGCCGTTCT TCAACAATA CCGTCCGCAG TTCTACTTCC 700
   GCACCACCGA CGTGACCGGT GTGGTGACGC TGCCGGAGGG CACCGAGATG 750
   GTGATGCCCG GTGACAACAC CAACATCTCG GTGAAGCTGA TCCAGCCCGT 800
55 CGCCATGGAC GACGGTSTGC GGTTC 825

```

## 2) INFORMATION FOR SEQ ID NO: 977

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 820 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-06

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 977

15	TATCCTTGTA	GTAGCTTCAA	CTGACGGACC	AATGCCACAA	ACTCGTGAGC	50
	ACATCCTTCT	TTCACGTCAG	GTTGGTGTTA	AACACCTTAT	CGTCTTCATG	100
	AACAAAGTTG	ACTTGGTTGA	CGACGAAGAA	TTGCTTGAAT	TGGTTGAAAT	150
	GGAAATCCGT	GACCTATTGT	CAGAATACGA	CTTCCCAGGT	GACGATCTTC	200
	CAGTTATCCA	AGGTTTCAGCA	CTTAAAGCTC	TTGAAGGTGA	CTCTAAATAC	250
20	GAAGACATCG	TTATGGAATT	GATGAACACA	GTTGATGAGT	ATATCCCAGA	300
	ACCAGAACGT	GACACTGACA	AACCATTGCT	TCTTCCAGTC	GAGGACGTAT	350
	TCTCAATCAC	TGGACGTGGT	ACAGTTGCTT	CAGGACGTAT	CGACCGTGGT	400
	ATCGTTAAAG	TCAACGACGA	AATCGAAATT	GTTGGTATCA	AAGAAGAAAC	450
	TCAAAAAGCA	GTTGTTACTG	GTGTTGAAAT	GTTCCTGAAA	CAACTTGACG	500
25	AAGGTCTTGC	TGGAGATAAC	GTAGGTGTCC	TTCTTCGTGG	TGTTCAACGT	550
	GATGAAATCG	AACGTGGACA	AGTTATCGCT	AAACCAGGTT	CAATCAACCC	600
	ACACACTAAA	TTCAAAGGTG	AAGTCTACAT	CCTTACTAAA	GAAGAAGGTG	650
	GACGTCACAC	TCCATTCTTC	AACAACCTACC	GTCCACAAAT	CTACTTCCGT	700
	ACTACTGACG	TTACAGGTTT	AATCGAACTT	CCAGCAGGTA	CTGAAATGGT	750
30	AATGCCTGGT	GATAACGTGA	CAATCGACGT	TGAGTTGATT	CACCCAATCG	800
	CCGTAGAACA	AGGTACTACA				820

## 35 2) INFORMATION FOR SEQ ID NO: 978

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 822 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium gordonae*  
 (B) STRAIN: M-Gor-1

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 978

50	GGCGCGATCC	TGGTGGTCGC	CGCCACCGAT	GGCCCGATGC	CGCAGACCCG	50
	TGAGCACGTG	CTGCTCGCCC	GTCAGGTGGG	CGTGCCCTAC	ATCCTGGTG	100
	CGCTGAACAA	GTCCGACGCG	GTCGACGACG	AGGAGCTGCT	CGAGCTCGTC	150
	GAGCTGGAGG	TCCGCGAGTT	GCTGGCCGCC	CAGGACTTCG	ACGAGGAAGC	200
55	TCCGGTGGTC	CGGGTCTCGG	CGCTGAAGGC	GCTCGAGGGC	GACGCCACCT	250
	GGGTGAAGTC	GGTAGAGGAC	TTGATGGACG	CGGTCGACGA	GTCGATTCCG	300
	GACCCGGTCC	GCGACACCGA	CAAGCCGTTC	CTGATGCCCG	TCGAGGACGT	350
	CTTCACCATC	ACCGGTCGTG	GCACCGTCGT	CACCGGCCGT	GTGGAGCGCG	400
	GCGTGGTGAA	CGTGAACGAG	GAAGTCGAGA	TCGTCGGCAT	CAAGCCGACC	450
60	AGCACCAAGA	CCACGGTCAC	CGGTGTGGAG	ATGTTCCGCA	AGCTGCTCGA	500

CCAGGGTCAG GCCGGTGACA ACGTCGGTCT GCTGCTGCGT GGTGTCAAGC 550  
 GTGAGGACGT CGAGCGCGGC CAGGTCGTCA TCAAGCCCGG CACCACCACT 600  
 CCGCACACCG AGTTCGAGGG TCAGGTCTAC ATCCTGTCCA AGGACGAGGG 650  
 CGGCCGGCAC ACGCCGTTCT TCAACAACTA CCGTCCGCAG TTCTACTTCC 700  
 5 GCACCACCGA CGTGACCGGT GTGGTGACGC TGCCGGAGGG CACCGAAATG 750  
 GTGATGCCCCG GTGACAACAC CAACATCTCG GTGAAGCTGA TCCAGCCCCG 800  
 CGCCATGGAC GACGGTCTGC GG 822

10

## 2) INFORMATION FOR SEQ ID NO: 979

## (i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 821 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

20

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-11

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 979

CTATCCTTGT AGTAGCTTCA ACTGACGGAC CAATGCCACA AACTCGTGAG 50  
 CACATCCTTC TTTCACGTCA GGTGTTGGTGT AAACACCTTA TCGTCTTCAT 100  
 GAACAAAGTT GACTTGGTTG ACGACGAAGA ATTGCTTGAA TTGGTTGAAA 150  
 30 TGGAAATCCG TGACCTATFG TCAGAATACG ACTTCCCAGG TGACGATCTT 200  
 CCAGTTATCC AAGGTTTCAGC ACTTAAAGCT CTTGAAGGTG ACTCTAAATA 250  
 CGAAGACATC GTTATGGAAT TGATGAACAC AGTTGATGAG TATATCCAG 300  
 AACCAGAACG TGACACTGAC AAACCATTGC TTCTTCCAGT CGAGGACGTA 350  
 TTCTCAATCA CTGGACGTGG TACAGTTGCT TCAGGACGTA TCGACCGTGG 400  
 35 TATCGTTAAA GTCAACGACG AAATCGAAAT CGTTGGTATC AAAGAAGAAA 450  
 CTCAAAAAGC AGTTGTTACT GGTGTTGAAA TGTTCGGTAA ACAACTTGAC 500  
 GAAGGTCTTG CTGGAGATAA CGTAGGTGTC CTTCTTCGTG GTGTTCAACG 550  
 TGATGAAATC GAACGTGGAC AAGTTATCGC TAAACCAGGT TCAATCAACC 600  
 CACACACTAA ATTCAAAGGT GAAGTCTACA TCCTTACTAA AGAAGAAGGT 650  
 40 GGACGTCACA CTCCATTCTT CAACAACCTAC CGTCCACAAT TCTACTTCCG 700  
 TACTACTGAC GTTACAGGTT CAATCGAACT TCCAGCAGGT ACTGAAATGG 750  
 TAATGCCTGG TGATAACGTG ACAATCGACG TTGAGTTGAT TCACCCAATC 800  
 GCCGTAGAAC AAGGTACTAC A 821

45

## 2) INFORMATION FOR SEQ ID NO: 980

## (i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 828 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 55 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Mycobacterium tuberculosis*  
 (B) STRAIN: ATCC 25177

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 980

```

5  GGTGCGATCC TGGTGGTTCG CGCCACCGAC GGCCCGATGC CCCAGACCCG      50
   CGAGCACGTT CTGCTGGCGC GTCAAGTGGG TGTGCCCTAC ATCCTGGTAG      100
   CGCTGAACAA GGCCGACGCA GTGGACGACG AGGAGCTGCT CGAACTCGTC      150
   GAGATGGAGG TCCGCGAGCT GCTGGCTGCC CAGGAATTCG ACGAGGACGC      200
   CCCGGTTGTG CGGGTCTCGG CGCTCAAGGC GCTCGAGGGT GACGCGAAGT      250
   GGGTTGCCTC TGTCGAGGAA CTGATGAACG CCGTCGACGA GTCGATTCCG      300
   GACCCGGTCC GCGAGACCGA CAAGCCGTTC CTGATGCCGG TCGAGGACGT      350
10  CTTACCATTT ACCGGCCGCG GAACCGTGGT CACCGGACGT GTGGAGCGCG      400
   GCGTGATCAA CGTGAACGAG GAAGTTGAGA TCGTCGGCAT TCGCCCATCG      450
   ACCACCAAGA CCACCGTCAC CGGTGTGGAG ATGTTCCGCA AGCTGCTCGA      500
   CCAGGGCCAG GCGGGCGACA ACGTTGGTTT GCTGCTGCGG GGCCTCAAGC      550
   GCGAGGACGT CGAGCGTGGC CAGGTTGTCA CCAAGCCCGG CACCACCACG      600
15  CCGCACACCG AGTTCGAAGG CAGGTCCTAC ATCCTGTCCA AGGACGAGGG      650
   CGGCCGGCAC ACGCCGTTCT TCAACAAC TA CCGTCCGACG TTCTACTTCC      700
   GCACCACCGA CGTGACCGGT GTGGTGACAC TGCCGGAGGG CACCGAGATG      750
   GTGATGCCCG GTGACAACAC CAACATCTCG GTGAAGTTGA TCCAGCCCGT      800
   CGCCATGGAC GAAGGTNTGC GTTTCGCG      828
20

```

## 2) INFORMATION FOR SEQ ID NO: 981

```

25  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 819 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear
30
   (ii) MOLECULE TYPE: Genomic DNA
   (vi) ORIGINAL SOURCE:
      (A) ORGANISM: Staphylococcus warneri
35  (B) STRAIN: CSG 144

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 981

```

40  CGGCGGTATC TTAGTAGTAT CTGCTGCAGA TGGCCCAATG CCACAAACTC      50
   GTGAACACAT CTTATTATCA CGTAACGTTG GTGTACCAGC TTTAGTTGTA      100
   TTCCTTAAACA AAGCTGACAT GGTGACGAC GAAGAATTAT TAGAATTAGT      150
   TGAAATGGAA GTTCGTGACT TATTATCTGA ATACGACTTC CCTGGTGACG      200
   ATGTACCAGT TATCGTTGGT TCTGCATTAA AAGCTTTAGA AGGCGACCCA      250
   GAATACGAAC AAAAAATCTT AGACTTAATG CAAGCTGTAG ATGACTACAT      300
45  CCCAACTCCA GAACGTGACT CTGATAAACC ATTCATGATG CCAGTTGAGG      350
   ACGTATTCTC AATCACTGGT CGTGGTACTG TAGCAACAGG CCGTGTGAA      400
   CGTGGTCAAA TCAAAGTCGG TGAAGAAGTT GAAATCATCG GTATCACTGA      450
   AGAAAGCAAG AAAACAACAG TTACAGGTGT AGAAATGTTC CGTAAATTAT      500
   TAGACTACGC TGAAGCTGGT GACAACATCG GTGCTTTATT ACGTGGTGTT      550
50  GCACGTGAAG ACGTACAACG TGGACAAGTA TTAGCAGCTC CTGGCTCTAT      600
   TACTCCACAC ACAAATTC AAGCTGATGT TTACGTTT TCTAAAGAAG      650
   AAGGTGGACG TCATACTCCA TTCTTACTA ACTACGCCC ACAATTCTAC      700
   TTCCGTACTA CTGACGTAAC TGGCGTTGTT CACTTACCAG AAGGTACTGA      750
   AATGGTTATG CCTGGCGATA ACGTAGAAAT GACTGTTGAA TTAATCGCTC      800
55  CAATCGCGAT TGAAGACG      819

```

## 2) INFORMATION FOR SEQ ID NO: 982

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 814 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
 (B) STRAIN: LSPQ 2583

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 982

```

15 CTATCCTTGT AGTAGCTTCA ACTGACGGAC CAATGCCACA AACTCGTGAG      50
   CACATCCTTC TTTCACGTCA GGTGTTGGT AAACACCTTA TCGTCTTCAT      100
   GAACAAGATC GACTTGTTTG ATGACGAAGA ATTGCTTGAA TTGGTTGAAA      150
   TGGAAATCCG TGACCTCTTG TCAGAATACG ACTTCCCAGG TGACGATCTT      200
   CCAGTTATCC AAGGTTTCAGC TCTTAAAGCT CTTGAAGGTG AACTAAGTA      250
20 CGAAGACATC ATCATGGAAT TGATGAACAC TGTTGATGAG TACATCCCAG      300
   AACCAGAACG TGATACTGAC AAACCTCTTC TTCTTCCAGT CGAAGACGTA      350
   TTCTCAATCA CTGGTCGTGG TACAGTTGCT TCAGGACGTA TCGACCGTGG      400
   TACTGTTTCGT GTCAACGATG AAATCGAAAT CGTTGGTATC AAAGAAGAAA      450
   TCCAAAAGC AGTTGTTACT GGTGTTGAAA TGTTCCGTAA ACAGCTTGAC      500
25 GAAGGTCCTG CAGGGGACAA CGTAGGTGTA CTTCTTCGTG GTATCCAACG      550
   TGATGAAATC GAACGTGGTC AAGTTATCGC TAAACCAGGT TCAATCAACC      600
   CACACACTAA ATTCAAGGGT GAAGTTTACA TCCTTACTAA AGAAGAAGGT      650
   GGACGTCACA CTCCATTCTT CAACAACACT CGTCCACAGT TCTACTTCCG      700
   TACAACGAC GTTACAGGTT CAATCGAACT TCCAGCAGGT ACTGAAATGG      750
30 TAATGCCTGG TGATAACGTA ACTATCGACG TTGAGTTGAT CCACCCAATC      800
   GCCGTTGAAC AAGG                                     814

```

## 35 2) INFORMATION FOR SEQ ID NO: 983

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: bases  
 (B) TYPE: 810 Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
 (B) STRAIN: ATCC 49456

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 983

```

50 CTATCCTTGT AGTAGCTTCA ACTGACGGAC CAATGCCACA AACTCGTGAG      50
   CACATCCTTC TTTCACGTCA GGTGTTGGT AAACACCTTA TCGTCTTCAT      100
   GAACAAAGTT GACTTGTTTG ACGACGAAGA ATTGCTTGAA TTGGTTGAAA      150
   TGGAAATCCG TGACCTATTG TCAGAATACG ACTTCCCAGG TGACGATCTT      200
55 CCAGTTATCC AAGGTTTCAGC TCTTAAAGCC CTTGAAGGTG AACTAATA      250
   CGAAGACATC GTTATGGAAT TGATGAACAC AGTTGATGAG TACATCCCAG      300
   AACCAGAACG TGACACTGAC AAACCATTCG TTCTTCCAGT CGAAGACGTA      350
   TTCTCAATCA CTGGTCGTGG TACAGTTGCT TCAGGACGTA TCGACCGTGG      400
   TATCGTTAAA GTCAACGACG AAATCGAAAT CGTTGGTATC AAAGAAGAAA      450
60 CTCAAAAGC AGTTGTTACT GGTGTTGAAA TGTTCCGTAA ACAACTTGAC      500

```

	GAAGGTCTTG	CCGGAGATAA	TGTAGGTGTC	CTTCTTCGTG	GTGTTCAACG	550
	TGATGAAATC	GAACGTGGAC	AAGTTATTGC	TAAACCAGGT	TCAATCAACC	600
	CACACACTAA	ATTCAAAGGT	GAAGTTTACA	TCCTTACTAA	AGAAGAAGGT	650
	GGACGTCACA	CTCCATTCTT	CAACAACACTAC	CGTCCACAAT	TCTACTTCCG	700
5	TACTACTGAC	GTTACAGGTT	CAATCGAACT	TCCAGCAGGT	ACTGAAATGG	750
	TAATGCCTGG	TGATAACGTG	ACAATCGACG	TTGAGTTGAT	CCACCCAATC	800
	GCCGTAGAAC					810

10

## 2) INFORMATION FOR SEQ ID NO: 984

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 817 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## 20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
(B) STRAIN: ATCC 903

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 984

	TCCTTGTAGT	AGCTTCAACT	GACGGACCAA	TGCCACAAAC	TCGTGAGCAC	50
	ATCCTTCTTT	CACGTCAGGT	TGGTGTTAAA	CACCTTATCG	TCTTCATGAA	100
	CAAGATCGAC	TTGGTTGATG	ACGAAGAATT	GCTTGAATTG	GTTGAAATGG	150
30	AAATCCGTGA	CCTCTTGTC	GAATACGACT	TCCCAGGTGA	CGATCTTCCA	200
	GTTATCCAAG	GTTTCAGCTCT	TAAAGCTCTT	GAAGGTGATA	CTAAGTACGA	250
	AGACATCATC	ATGGAATTGA	TGAACACTGT	TGATGAGTAC	ATCCCAGAAC	300
	CAGAACGTGA	TACTGACAAA	CCTCTTCTTC	TTCCAGTCGA	AGACGTATTC	350
	TCAATCACTG	GTCGTGGTAC	AGTTGCTTCA	GGACGTATCG	ACCGTGGTAC	400
35	TGTTTCGTGTC	AACGATGAAA	TCGAAATCGT	TGGTATCAAA	GAAGAAATCC	450
	AAAAAGCAGT	TGTTACTGGT	GTTGAAATGT	TCCGTAAACA	GCTTGACGAA	500
	GGTCTTGCA	GGGACAACGT	AGGTGTACTT	CTTCGTGGTA	TCCAACGTGA	550
	TGAAATCGAA	CGTGGTCAAG	TTATCGCTAA	ACCAGGTTCA	ATCAACCCAC	600
	ACACTAAAT	CAAGGGTGAA	GTTTACATCC	TTACTAAAGA	AGAAGGTGGA	650
40	CGTCACACTC	CATTCTTCAA	CAACTACCGT	CCACAGTTCT	ACTTCCGTAC	700
	AACTGACGTT	ACAGGTTCAA	TCGAACTTCC	AGCAGGTACT	GAAATGGTAA	750
	TGCCTGGTGA	TAACGTAACT	ATCGACGTTG	AGTTGATCCA	CCCAATCGCC	800
	GTTGAACAAG	GTACTAC				817

45

## 2) INFORMATION FOR SEQ ID NO: 985

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 798 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## 55 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus oralis*  
(B) STRAIN: ATCC 35037

60

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 985

	GACGGACCAA	TGCCACAAAC	TCGTGAGCAC	ATCCTTCTTT	CACGTCAGGT	50
	TGGTGTAA	CACCTTATTG	TCTTCATGAA	CAAAATTGAC	TTGGTAGACG	100
5	ACGAAGAATT	GCTTGAATTG	GTTGAAATGG	AAATCCGTGA	CCTCTTGTCA	150
	GAATACGACT	TCCCAGGTGA	CGATCTTCCA	GTTATCCAAG	GTTTCAGCTCT	200
	TAAAGCTCTT	GAAGGTGACT	CTAAATACGA	AGACATCATT	ATGGAATTGA	250
	TGAACACTGT	TGATGAGTAC	ATCCCAGAAC	CAGAACGTGA	CACTGAAAAA	300
	CCATTGCTTC	TTCCAGTCGA	AGACGTATTC	TCAATCACTG	GACGTGGTAC	350
10	AGTTGCTTCA	GGACGTATCG	ACCGTGGTAC	TGTTTCGTGTC	AACGACGAAA	400
	TCGAAATCGT	TGGTATCAAA	GAAGAAACTC	AAAAAGCAGT	TGTTACTGGT	450
	GTTGAAATGT	TCCGTAAACA	ACTTGACGAA	GGTCTTGCCG	GAGATAACGT	500
	AGGTGTCTT	CTTCGTGGTG	TTCAACGTGA	CGAAATCGAA	CGTGGACAAG	550
	TTATCGCTAA	ACCAGGTTCA	ATCAACCCAC	ACACTAAATT	TAAAGGTGAA	600
15	GTCTACATCC	TTACTAAAGA	AGAAGGTGGA	CGTCACACTC	CATTCTTCAA	650
	CAACTACCGT	CCACAATTCT	ACTTCCGTAC	TACTGACGTT	ACAGGTTCAA	700
	TCGAACTTCC	TGCAGGTACT	GAAATGGTAA	TGCCTGGTGA	TAACGTGACT	750
	ATCGACGTTG	AGTTGATCCA	CCCAATCGCC	GTAGAACAAG	GTACTACA	798

20

## 2) INFORMATION FOR SEQ ID NO: 986

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 815 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 30 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-05

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 986

	TTGTAGTAGC	TTCAACTGAC	GGACCAATGC	CACAAACTCG	TGAGCACATC	50
	CTTCTTTTAC	GTCAGGTTGG	TGTTAAACAC	CTTATCGTCT	TCATGAACAA	100
40	AGTTGACTTG	GTTGACGACG	AAGAATTGCT	TGAATTGGTT	GAAATGGAAA	150
	TCCGTGACCT	ATTGTCAGAA	TACGACTTCC	CAGGTGACGA	TCTTCCAGTT	200
	ATCCAAGGTT	CAGCACTTAA	AGCTCTTGAA	GGTGACTCTA	AATACGAAGA	250
	CATCGTTATG	GAATTGATGA	ACACAGTTGA	TGAGTATATT	CCAGAACCAG	300
	AACGTGACAC	TGACAAACCA	TTGCTTCTTC	CAGTCGAGGA	CGTATTCTCA	350
45	ATCACTGGAC	GTGGTACAGT	TGCTTCAGGA	CGTATCGACC	GTGGTATCGT	400
	TAAAGTCAAC	GACGAAATCG	AAATCGTTGG	TATCAAAGAA	GAAACTCAA	450
	AAGCAGTTGT	TACTGGTGTT	GAAATGTTCC	GTAAACAAC	TGACGAAGGT	500
	CTTGCTGGAG	ATAACGTAGG	TGTCCTTCTT	CGTGGTGTTT	AACGTGATGA	550
	AATCGAACGT	GGACAAGTTA	TCGCTAAACC	AGGTTCAATC	AACCCACACA	600
50	CTAAATTCAA	AGGTGAAGTC	TACATCCTTA	CTAAAGAAGA	AGGTGGACGT	650
	CACACTCCAT	TCTTCAACAA	CTACCGTCCA	CAATTCTACT	TCCGTACTAC	700
	TGACGTTACA	GGTTCAATCG	AACTTCCAGC	AGGTACTGAA	ATGGTAATGC	750
	CTGGTGATAA	CGTGACAATC	GACGTTGAGT	TGATTCAACC	AATCGCCGTA	800
	GAACAAGGTA	CTACA				815

55

## 2) INFORMATION FOR SEQ ID NO: 987

## 60 (i) SEQUENCE CHARACTERISTICS:

752

(A) LENGTH: 832 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus hirae*  
 (B) STRAIN: ATCC 8043

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 987

15	CGGACAATCT	TGGTTGTTTC	TGCTACGGAT	GGCCCAATGC	CTCAAACCTCG	50
	TGAACACATT	TTACTTTCTC	GCCAAGTAGG	CGTGAAATAT	TTGATTGTTT	100
	TCTTGAACAA	AACAGATTTA	GTCGATGATG	AAGAATTAAT	TGATCTAGTA	150
	GAAATGGAAG	TTCGTGAACT	ATTAAGCGAA	TATGGATTTC	CAGGTGATGA	200
	TACACCTGTT	ATCAAAGGAT	CAGCATTAAA	AGCATTACAA	GGTGATCCTG	250
	ATGCAGAAGC	AGCTATCATG	GAATTGATGG	ACACAGTCGA	TGAATACATC	300
20	CCAACACCAG	AACGTGATAC	GGACAAACCA	TTATTGTTAC	CCGCCGAAGA	350
	TGTATTTTCA	ATTACTGGAC	GTGGAACAGT	AGCATCTGGT	CGGATTGACC	400
	GTGGGGCTGT	TAGAGTCGGT	GACGAAATCG	AAATCGTAGG	GATCAAACCA	450
	GAAACGCAAA	GAGCCGTAGT	AACAGGAGTT	GAAATGTTCC	GCAAAAACGCT	500
	TGATTACGGT	GAAGCAGGGG	ATAACGTAGG	TGTGTTATTA	CGTGGGATTC	550
25	AAAGAGAAGA	CATCGAACGT	GGCCAAGTGA	TTGCCAAACC	TGGTTC AATT	600
	ACACCTCATA	CTAAATTCAA	AGCAGAAGTT	TAGGTTTGA	CTAAAGAAGA	650
	AGGCGGACGT	CATACACCAT	TCTTCAATAA	TTATCGACCA	CAATTTTATT	700
	TCCGCACAAC	TGACGTAAAC	GGAACAATTG	TTTTGCCAGA	AGGAACGGAA	750
	ATGGTCATGC	CTGGCGACAA	CGTAACGATC	GATGTAGAAT	TGATTCATCC	800
30	TGTTGCTATT	GAAAACGGGA	CGACTTTCTC	CA		832

2) INFORMATION FOR SEQ ID NO: 988

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 835 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus mundtii*  
 (B) STRAIN: ATCC 43186

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 988

50	TGGTGCGATC	TTAGTTGTTT	CTGCGACAGA	TGGACCAATG	CCTCAAACCTC	50
	GCGAGCATAT	CCTTTTATCA	CGTCAAGTTG	GTGTAAAATA	TTTGATTGTA	100
	TTTTTGAATA	AAGTTGATTT	GGTCGATGAT	GAAGAATTGA	TCGATCTTGT	150
	AGAAATGGAA	GTTCTGTGAAT	TACTGAATGA	ATATGGTTTC	CCAGGTGATG	200
	ACACACCTGT	CATCAAAGGC	TCTGCATTGA	AAGCATTGCA	AGGTGACCCA	250
55	GAAGCAGAAG	CTGCGATCAA	TGAATTGATG	GAAACAGTGG	ATGACTATAT	300
	CCCAACACCA	GAACGTGATA	CCGACAAACC	ATTGCTTTTA	CCAGTTGAAG	350
	ATGTTTTTCTC	AATCACTGGT	CGCGGAACGG	TAGCATCTGG	TCGTATCGAC	400
	CGTGGAGCCG	TTCTGTCTCG	TGATGAAATC	GAGATCATCG	GAATCAAACC	450
	TGAAACGAAA	AAAGCGGTTG	TGACAGGGGT	AGAGATGTTC	CGTAAAACGT	500
60	TAGATTATGG	CGAAGCTGGA	GATAACGTAG	GAATCTTGTT	ACGTGGTATC	550



	CAAAGAGAAG	ATATTGAACG	TGGACAAGTA	ATTGCGAAAC	CTGGTTCAAT	600
	CACACCACAT	ACAAAATTCA	AAGCGGAAGT	TTATGTATTG	ACGAAAGAAG	650
	AAGGCGGACG	TCATACACCA	TTCTTCAATA	ACTACGCCC	ACAATTTTAT	700
	TTCCGCACAA	CAGATGTAAC	AGGTACGATC	GTGTTGCCAG	AAGGAACAGA	750
5	AATGGTCATG	CCTGGAGACA	ACGTAACCAT	CGAGGTAGAG	TTGATCCATC	800
	CAGTGGCAAT	CGAACAAGGA	ACGACTTTCT	CTATT		835

## 10 2) INFORMATION FOR SEQ ID NO: 989

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 832 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 20 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus raffinosus*  
 (B) STRAIN: ATCC 49427

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 989

25	TGCGATCTTA	GTA GTATCTA	CTACTGATGG	ACCAATGCCT	CAAACACGTG	50
	AGCACATTTT	GTTATCACGC	CAAGTAGGTG	TTAAATACTT	GATCGTCTTC	100
	TTGAACAAAG	TTGATTTAGT	CGACGATGAA	GAATTGATTG	ATTTAGTTGA	150
	AATGGAAGTA	CGTGAGTTAC	TTTCAGAATA	TGGTTTCCCA	GGCGATGATA	200
30	TTCTGTCTCT	TAAAGGTTCA	GCTCTGAAAG	CTTTAGAAGG	CGATCCTGAA	250
	CAAGAACAAG	TAATCATGGA	CCTAATGGAT	ACGGTTGACG	AATATATCCC	300
	AACACCAGAA	CGTGACACTG	ACAAACCATT	CTTGTTACCA	GTGGAAGATG	350
	TTTTCTCAAT	CACAGGACGT	GGGACTGTTG	CATCTGGTCG	TATTGATCGT	400
	GGGGAAGTTA	AAGTCGGTGA	CGAAGTTGAA	ATTATCGGGA	TCAAACCTGA	450
35	AGTTCAAAAG	GCTGTGCTAA	CTGGACTTGA	AATGTTCCGT	AAAACATTGG	500
	ATTATGGTGA	AGCTGGAGAT	AACGTTGGGG	TTCTATTACG	TGGTATTACA	550
	CGTGATGAAA	TCGAACGTGG	TCAAGTATTA	GCTAAACCAG	GTTCAATTAC	600
	ACCACATACG	AAATTCAGTG	CAGAAGTTTA	TGTGTTGACG	AAAGAAGAAG	650
	GTGGACGTCA	TACACCATTC	TTTAACAAC	ATCGTCCTCA	ATTCTACTTC	700
40	CGTACAACAG	ACGTTACCGG	TAATATCGTG	TTGCCAGAAG	GTACTGAAAT	750
	GGTCATGCCT	GGCGATAACG	TAACAATCGA	CGTTGAATTA	ATCCATCCAA	800
	TCGCCGTAGA	AAAAGGAACA	ACTTTCTCTA	TT		832

## 45 2) INFORMATION FOR SEQ ID NO: 990

## (i) SEQUENCE CHARACTERISTICS:

- 50 (A) LENGTH: 154 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 55 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bacillus anthracis*  
 (B) STRAIN: CIP 9440

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 990

	GGATCCTGTA	TATGCACAAA	AACTAGGTGT	TAACATCGAT	GAATTACTAT	50
	TATCACAACC	TGATACAGGG	GAGCAAGGTT	TAGAAATCGC	AGAAGCACTT	100
	GTACGAAGTG	GTGCGGTGA	TATTATCGTA	ATTGACTCTG	TAGCAGCTCT	150
5	TGTA					154

## 2) INFORMATION FOR SEQ ID NO: 991

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 384 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Prevotella melaninogenica*  
 (B) STRAIN: ATCC 25845

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 991

25	GCCATTGCAG	AGGCACAGAA	GCAGGGCGGT	ATTGCAGCCT	TCATTGATGC	50
	TGAGCACGCC	TTCGACCGTT	TCTATGCAGA	GAAGTTAGGT	GTGGATGTTG	100
	ATAACCTTTG	GGTTTCACAG	CCAGACAATG	GTGAGCAGGC	TTTAGAGATT	150
	GCCGACCAGC	TGATTCGCTC	TTCCGCTATT	GACATTCTCG	TTGTCCGACTC	200
	AGTTGCAGCC	TTGACTCCAA	AGAAGGAGAT	TGAGGGTGAC	ATGGGTGACT	250
30	CTGCAGTAGG	TTTACAAGCA	CGACTGATGA	GTCAGGCATT	GCGTAAACTT	300
	ACCTCAACAA	TCGCAAAAAC	TAATACTTGC	TGCATCTTCA	TCAACCAGTT	350
	GCGTGAGAAG	ATTGGTGTGA	TGTTTGGTAA	TCCA		384

## 2) INFORMATION FOR SEQ ID NO: 992

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 624 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus casseliflavus*  
 (B) STRAIN: R760

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 992

	ACACGTGAAC	ACATCTTGTT	ATCACGTAAC	GTTGGTGTAC	CATACATCGT	50
	TGTTTTCTTA	AACAAAATGG	ATATGGTTGA	TGACGAAGAA	TTACTAGAAT	100
	TAGTTGAAAT	GGAAGTTCGT	GACTTATTGT	CAGAATATGA	CTTCCCAGGC	150
55	GACGATGTTT	CTGTAATCGC	TGGTTCTGCT	TTGAAAGCTC	TTGAAGGCGA	200
	TGCTTCATAC	GAAGAAAAAA	TCATGGAATT	AATGGCTGCA	GTTGACGAAT	250
	ACGTTCCAAC	TCCAGAACGT	GACACTGACA	AACCATTTCAT	GATGCCAGTC	300
	GAAGACGTAT	TCTCAATCAC	TGGACGTGGT	ACTGTTGCTA	CAGGCCGTGT	350
	TGAACGTGGA	CAAGTTCGCG	TTGGTGACGA	AGTTGAAATC	GTTGGTATTG	400
60	CTGAAGAAAC	TGCTAAAACA	ACTGTAAC TG	GTGTTGAAAT	GTTCCGTAAA	450

TTGTTAGACT	ATGCTGAAGC	AGGGGATAAC	ATTGGTGCAT	TGCTACGTGG	500
TGTTGCTCGT	GAAGACATCC	AACGTGGACA	AGTATTGGCT	AAAGCTGGTA	550
CAATCACACC	TCATACAAAA	TTTAAAGCTG	AAGTTTACGT	TTTAACAAAA	600
GAAGAAGGTG	GACGTCATAC	ACCA			624

5

## 2) INFORMATION FOR SEQ ID NO: 993

10 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 756 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

20 (A) ORGANISM: *Streptococcus pyogenes*  
 (C) ACCESSION NUMBER: extracted from U40453

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 993

25	ATGGAAAACA	ATAAAAAAGT	ATTGAAGAAA	ATGGTATTTT	TTGTTTtagT	50
	GACATTTCTT	GGACTAACAA	TCTCGCAAGA	GGTATTTGCT	CAACAAGACC	100
	CCGATCCAAG	CCAAC TTCAC	AGATCTAGTT	TAGTTAAAAA	CCTTCAAAAT	150
	ATATATTTTC	TTTATGAGGG	TGACCCTGTT	ACTCACGAGA	ATGTGAAATC	200
	TGTTGATCAA	CTTTTATCTC	ACGATTTAAT	ATATAATGTT	TCAGGGCCAA	250
	ATTATGATAA	ATTAAAAACT	GAAC TTAAGA	ACCAAGAGAT	GGCAACTTTA	300
30	TTTAAGGATA	AAAACGTTGA	TATTTATGGT	GTAGAATATT	ACCATCTCTG	350
	TTATTTATGT	GAAAAATGCAG	AAAGGAGTGC	ATGTATCTAC	GGAGGGGTAA	400
	CAAATCATGA	AGGGAATCAT	TTAGAAATTC	CTAAAAAGAT	AGTCGT TAAA	450
	GTATCAATCG	ATGGTATCCA	AAGCCTATCA	TTTGATATTG	AAACAAATAA	500
	AAAAATGGTA	ACTGCTCAAG	AATTAGACTA	TAAAGTTAGA	AAATATCTTA	550
35	CAGATAATAA	GCAACTATAT	ACTAATGGAC	CTTCTAAATA	TGAAACTGGA	600
	TATATAAAGT	TCATACCTAA	GAATAAAGAA	AGTTTTTGGT	TTGATTTTTT	650
	CCCTGAACCA	GAATTTACTC	AATCTAAATA	TCTTATGATA	TATAAAGATA	700
	ATGAAACGCT	TGACTCAAAC	ACAAGCCAAA	TTGAAGTCTA	CCTAACAACC	750
	AAGTAA					756

40

## 2) INFORMATION FOR SEQ ID NO: 994

45 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 994

55 TGGACTAACA ATCTCGCAAG AGG

23

## 2) INFORMATION FOR SEQ ID NO: 995

60

756

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 995

ACATTCTCGT GAGTAACAGG GT

22

2) INFORMATION FOR SEQ ID NO: 996

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 996

ACAAATCATG AAGGGAATCA TTTAG

25

2) INFORMATION FOR SEQ ID NO: 997

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 997

CTAATTCTTG AGCAGTTACC ATT

23

2) INFORMATION FOR SEQ ID NO: 998

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 998

GGAGGGGTAA CAAATCATGA AGG

23

## 2) INFORMATION FOR SEQ ID NO: 999

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 999  
15 TTGACCTTGT TGATGACGAA GAG

23

## 2) INFORMATION FOR SEQ ID NO: 1000

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1000  
30 TTAGTGTGTG GGTGATTGA ACT

23

## 35 2) INFORMATION FOR SEQ ID NO: 1001

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1001  
AAGAGTTGCT TGAATTAGTT GAG

23

## 50 2) INFORMATION FOR SEQ ID NO: 1002

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 894 bases  
55 (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: Genomic DNA  
60

758

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pyogenes*  
 (B) STRAIN: ATCC 700294

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1002

```

AACATGATCA CTGGTGCCGC TCAAATGGAC GGAGCTATCC TTGTAGTTGC      50
TTCAACTGAT GGACCAATGC CACAAACTCG TGAGCACATC CTTCTTTCAC      100
GTCAGGTTGG TGTTAAACAC CTTATCGTGT TCATGAACAA AGTTGACCTT      150
10 GTTGATGACG AAGAGTTGCT TGAATTAGTT GAGATGGAAA TTCGTGACCT      200
TCTTTCAGAA TACGATTTC CAGGTGATGA CCTTCCAGTT ATCCAAGGTT      250
CAGCTCTTAA AGCTCTTGAA GGCGACACTA AATTTGAAGA CATCATCATG      300
GAATTGATGG ATACTGTTGA TTCATACATT CCAGAACCAG AACGCGACAC      350
TGACAAACCA TTGCTTCTTC CAGTCGAAGA CGTATTCTCA ATTACAGGTC      400
15 GTGGTACAGT TGCTTCAGGA CGTATCGACC GTGGTACTGT TCGTGTCAAC      450
GACGAAATCG AAATCGTTGG TATCAAAGAA GAAACTAAAA AAGCTGTTGT      500
TACTGGTGTG GAAATGTTCC GTAAACAAC TGAAGAAGGT CTTGCAGGAG      550
ACAACGTAGG TATCCTTCTT CGTGGTGTTC AACGTGACGA AATCGAACGT      600
GGTCAAGTTA TTGCTAAACC AAGTTCAATC AACCACACA CTAAATTCAA      650
20 AGGTGAAGTA TATATCCTTT CTAAAGACGA AGGTGGACGT CACTCCAT      700
TCTTCAACAA CTACCGTCCA CAATTCTACT TCCGTACAAC TGACGTAACA      750
GGTTCAATCG AACTTCCAGC AAGTACAGAA ATGGTTATGC CTGGTGATAA      800
CGTGACAATC AACGTTGAGT TGATCCACCC AATCGCCGTA GAACAAGGTA      850
CTACTTTCTC AATCCGTGAA GGTGGACGTA CTGTTGGTTC AGGT          894
25

```

## 2) INFORMATION FOR SEQ ID NO: 1003

## 30 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 332 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Bacillus cereus*  
 (B) STRAIN: ATCC 14579

## 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1003

```

AATGGATCCT GTATATGCAC AAAAATTAGG CGTTAACATA GATGAATTAC      50
45 TATTATCACA GCCTGATACA GGGGAGCAAG GATTAGAAAT CGCGGAAGCA      100
CTTGTACGAA GTGGTGCGGT TGACATTATC GTAATTGACT CTGTAGCAGC      150
TCTTGTACCG AAAGCAGAGA TTGAAGGCGA CATGGGTGAC TCACACGTAG      200
GTTTACAAGC ACGTTTAATG TCACAAGCAC TTCGTAAGCT TTCAGGAGCA      250
ATCAACAAAT CAAAAACAAT TGCAATCTTT ATTAACCAAA TTCGTGAAAA      300
50 AGTTGGGGTT ATGTTCGGAA ACCCAGAAAC AA          332

```

## 2) INFORMATION FOR SEQ ID NO: 1004

## 55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1212 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-01

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1004

10	ACCAAGAAGC	TCAAAAACAT	CTGTGGGATA	TCTACAACCTC	CGATCAATAC	50
	GTCTCTTACC	CTGACGATGA	TTTGCAAGTC	GCATCTACGG	TCGTAGATGT	100
	TTCAAATGGT	AAAGTCATCG	CCCAACTTGG	AGCTCGTCAC	CAAGCAAGTA	150
	ACGTTTCATT	TGGTACCAAC	CAAGCTGTGG	AAACCAATCG	TGACTGGGGT	200
	TCTGCTATGA	AACCAATCAC	CGATTATGCA	CCTGCCATAG	AATACGGTGT	250
15	TTATGATTCC	ACTGCAACTA	TGGTTAATGA	TATTCCTTAT	AACTATCCGG	300
	GAACAAGCAC	ACCTGTCTAC	AACTGGGATA	GAGCATATTT	CGGTAATATT	350
	ACTCTGCAAT	ATGCTCTTCA	ACAATCACGA	AATGTCACAG	CCGTTGAGAC	400
	TTTGAATAAG	GTCGGTCTAG	ATAGAGCTAA	AACCTTCCTT	AATGGTCTTG	450
	GTATCGACTA	TCCAAGCATG	CATTATGCAA	ACGCCATTTC	AAGTAATACA	500
20	ACAGAATCTA	ATAAACAATA	CGGAGCAAGT	AGTGAAAAAA	TGGCTGCTGC	550
	TTATGCTGCC	TTTGCAAATG	GTGGCACTTA	CTATAAACCA	ATGTATATCC	600
	ATAAAGTCGT	CTTCAGTGAT	GGAAGTAAAA	AAGAGTTCTC	TAATGTCGGA	650
	ACTCGTGCCA	TGAAGGAAAC	GACAGCCTAT	ATGATGACCG	ACATGATGAA	700
	AACAGTCTTG	ACTTATGGAA	CTGGGCGTGG	AGCCTATCTT	CCTTGGCTTC	750
25	CTCAAGCTGG	TAAAACAGGA	ACCTCTAACT	ATACAGATGA	GGAAGTTGAA	800
	AACCACATCA	AGAACACTGG	CTATGTAGCT	CCAGATGAAA	TGTTTGTGG	850
	TTATACTCGT	AAGTATTCTA	TGGCTGTATG	GACAGGTTAT	TCGAATCGTT	900
	TAACTCCTAT	CGTTGGAGAT	GGTTTCCTAG	TTGCAGCTAA	AGTTTATCGC	950
	TCAATGATAA	CGTATCTATC	AGAAGATACT	CATCCAGAAG	ACTGGACGAT	1000
30	GCCAGACGGA	CTTTTCAGAA	ACGGGGAATT	TGTATTCAAA	AATGGAGCTC	1050
	GCCCAATATG	GACTGAACCC	TCTACTCAAC	AATCCTCAAC	AGCTGAAAGT	1100
	TCAAGCTCAT	CATCAGATAG	TTCAACTTCA	CAGTCTAGCT	CAACCACTCC	1150
	AAGCACAAAT	AATAGTACGA	CTACCAATCC	TAACAATAAT	ACGCAACAAT	1200
	CAAATACAAC	CC				1212
35						

## 2) INFORMATION FOR SEQ ID NO: 1005

40 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1212 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

45 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 50 (A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-02

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1005

55	AAGCTCAAAA	ACATCTGTGG	GATATCTACA	ACTCCGATCA	ATACGTCTCT	50
	TACCCTGACG	ATGATTGCA	AGTCGCATCT	ACGGTCGTAG	ATGTTTCAAA	100
	TGGTAAAGTC	ATCGCCCAAC	TTGGAGCTCG	TCACCAAGCA	AGTAACGTTT	150
	CATTTGGTAC	CAACCAAGCT	GTGGAAACCA	ATCGTGACTG	GGGTTCTGCT	200
	ATGAAACCAA	TCACCGATTA	TGCACCTGCC	ATAGAATACG	GTGTTTATGA	250
	TTCCACTGCA	ACTATGGTTA	ATGATATTCC	TTATAACTAT	CCGGGAACAA	300
60	GCACACCTGT	CTACAACTGG	GATAGAGCAT	ATTTCCGGTAA	TATTACTCTG	350

	CAATATGCTC	TTCAACAATC	ACGAAATGTC	ACAGCCGTTG	AGACTTTGAA	400
	TAAGGTCGGT	CTAGATAGAG	CTAAAACCTT	CCTTAATGGT	CTTGGTATCG	450
	ACTATCCAAG	CATGCATTAT	GCAAACGCCA	TTTCAAGTAA	TACAACAGAA	500
	TCTAATAAAC	AATACGGAGC	AAGTAGTGAA	AAAATGGCTG	CTGCTTATGC	550
5	TGCCTTTGCC	AACGGTGGAA	TTTACCACAA	ACCCATGTAT	ATCAATAAGG	600
	TCGTCTTCAG	TGACGGTAGT	AAAAAAGAAT	TTTCAGATGT	AGGTACACGA	650
	GCTATGAAAG	AAACAACCTG	TTACATGATG	ACCGAAATGA	TGAAAACCTGT	700
	CTTGGCATAAC	GGAACCTGGT	GTGGAGCCTA	TCTCCCATGG	TTAGCGCAAG	750
	CTGGTAAGAC	AGGTACTTCT	AACTACACAG	ATGATGAAAT	TGAAAAACAC	800
10	ATCAAGAACA	CTGGCTATGT	AGCTCCAGAT	GAAATGTTTG	TTGGTTATAC	850
	TCGTAAGTAT	TCTATGGCTG	TATGGACAGG	TTATTCGAAT	CGTTTAACTC	900
	CTATCGTTGG	AGATGGTTTC	CTAGTTGCAG	CTAAAGTTTA	TCGCTCAATG	950
	ATAACGTATC	TGTCTGAAGG	AAGCAATCCA	GAGGATTGGA	ATATACCAGA	1000
	GGGGCTCTAC	AGAAATGGAG	AATTCGTATT	TAAAAATGGT	GCTCGTTCTA	1050
15	CGTGGAGCTC	ACCTGCTCCA	CAACAACCCC	CATCAACTGA	AAGTTCAAGC	1100
	TCATCATCAG	ATAGTTCAAC	TTCACAGTCT	AACTCAACCA	CTCCAAGCAC	1150
	AAATAATAGT	ACGACTACCA	ATCCTAACAA	TAATACGCAA	CAATCAAATA	1200
	CAACCCCTGA	TC				1212

20

## 2) INFORMATION FOR SEQ ID NO: 1006

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1213 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-03
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1006

	ACCAAGAAGC	TCAAAAACAT	CTGTGGGATA	TCTACAACCTC	CGATCAATAC	50
	GTCTCTTACC	CTGACGATGA	TTTGCAAGTC	GCATCTACGG	TCGTAGATGT	100
40	TTCAAATGGT	AAAGTCATCG	CCCAACTTGG	AGCTCGTCAC	CAAGCAAGTA	150
	ACGTTTCATT	TGGTACCAAC	CAAGCTGTGG	AAACCAATCG	TGACTGGGGT	200
	TCTGCTATGA	AACCAATCAC	CGATTATGCA	CCTGCCATAG	AATACGGTGT	250
	TTATGATTCC	ACTGCAACTA	TGGTTAATGA	TATTCCTTAT	AACTATCCGG	300
	GAACAAGCAC	ACCTGTCTAC	AACTGGGATA	GAGCATATTT	CGGTAATATT	350
45	ACTCTGCAAT	ATGCTCTTCA	ACAATCACGA	AATGTCACAG	CCGTTGAGAC	400
	TTTGAATAAG	GTCGGTCTAG	ATAGAGCTAA	AACCTTCCTT	AATGGTCTTG	450
	GTATCGACTA	TCCAAGCATG	CATTATGCAA	ACGCCATTTT	AAGTAATACA	500
	ACAGAATCTA	ATAAACAATA	CGGAGCAAGT	AGTGAAAAAA	TGGCTGCTGC	550
	TTATGCTGCG	TTTGCCAACG	GTGGAATTTA	CCACAAACCC	ATGTATATCA	600
50	ATAAGGTCGT	CTTCAGTGAC	GGTAGTAAAA	AAGAATTTTC	AGATGTAGGT	650
	ACACGAGCTA	TGAAAGAAAC	AACTGCTTAC	ATGATGACCG	AAATGATGAA	700
	AACTGTCTTG	GCATACGGAA	CTGGTCGTGG	AGCCTATCTC	CCATGGTTAG	750
	CGCAAGCTGG	TAAGACAGGT	ACTTCTAACT	ACACAGATGA	TGAAATTGAA	800
	AAACACATCA	AGAACACTGG	CTATGTAGCT	CCAGATGAAA	TGTTTGTGTT	850
55	TTTACTCGT	AAGTATTCTA	TGGCTGTATG	GACAGGTTAT	TCGAATCGTT	900
	TAACTCCTAT	CGTTGGAGAT	GGTTTCCTAG	TTGCAGCTAA	AGTTTATCGC	950
	TCAATGATAA	CGTATCTGTC	TGAAGGAAGC	AATCCAGAGG	ATTGGAATAT	1000
	ACCAGAGGGG	CTCTACAGAA	ATGGAGAATT	CGTATTTAAA	AATGGTGCTC	1050
	GTTCTACGTG	GAGCTCACCT	GCTCCACAAC	AACCCCATC	AACTGAAAGT	1100
60	TCAAGCTCAT	CATCAGATAG	TTCAACTTCA	CAGTCTAACT	CAACCACTCC	1150



AAGCACAAAT AATAGTACGA CTACCAATCC TAACAATAAT ACGCAACAAT 1200  
 CAAATACAAC CCC 1213

5

## 2) INFORMATION FOR SEQ ID NO: 1007

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1218 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

15

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-04

## 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1007

	ACCAAGAAGC	TCAAAAACAT	CTGTGGGATA	TCTACAAC	CGATCAATAC	50
	GTCTCTTACC	CTGACGATGA	TTTGCAAGTC	GCATCTACGG	TCGTAGATGT	100
	TTCAAATGGT	AAAGTCATCG	CCCAACTTGG	AGCTCGTCAC	CAAGCAAGTA	150
25	ACGTTTCATT	TGGTACCAAC	CAAGCTGTGG	AAACCAATCG	TGACTGGGGT	200
	TCTGCTATGA	AACCAATCAC	CGATTATGCA	CCTGCCATAG	AATACGGTGT	250
	TTATGATTCC	ACTGCAACTA	TGGTTAATGA	TATTCCTTAT	AACTATCCGG	300
	GAACAAGCAC	ACCTGTCTAC	AACTGGGATA	GAGCATATTT	CGGTAATATT	350
	ACTCTGCAAT	ATGCTCTTCA	ACAATCACGA	AATGTCACAG	CCGTTGAGAC	400
30	TTTGAATAAG	GTCCGGTCTAG	ATAGAGCTAA	AACCTTCCTT	AATGGTCTTG	450
	GTATCGACTA	TCCAAGCATG	CATTATGCAA	ACGCCATTTC	AAGTAATACA	500
	ACAGAATCTA	ATAAAACAATA	CGGAGCAAGT	AGTGAAAAAA	TGGCTGCTGC	550
	TTATGCTGCC	TTTGCAAATG	GTGGCACTTA	CTATAAACCA	ATGTATATCC	600
	ATAAAGTCGT	CTTCAGTGAT	GGAAGTAAAA	AAGAGTTCTC	TAATGTCGGA	650
35	ACTCGTGCCA	TGAAGGAAAC	GACAGCCTAT	ATGATGACCG	ACATGATGAA	700
	AACAGTCTTG	ACTTATGGAA	CTGGGCGTGG	AGCCTATCTT	CCTTGGCTTC	750
	CTCAAGCTGG	TAAACACAGGA	ACCTCTAACT	ATACAGATGA	GGAAGTTGAA	800
	AACCACATCA	AGAACACTGG	CTATGTAGCT	CCAGATGAAA	TGTTTGTGTTG	850
	TTATACTCGT	AAGTATTCTA	TGGCTGTATG	GACAGGTTAT	TCGAATCGTT	900
40	TAACTCCTAT	CGTTGGAGAT	GGTTTCCTAG	TTGCAGCTAA	AGTTTATCGC	950
	TCAATGATAA	CGTATCTATC	AGAAGATACT	CATCCAGAAG	ACTGGACGAT	1000
	GCCAGACGGA	CTTTTCAGAA	ACGGGGAATT	TGTATTCAAA	AATGGAGCTC	1050
	GCCCAATATG	GACTGAACCC	TCTACTCAAC	AATCCTCAAC	AGCTGAAAGT	1100
	TCAAGCTCAT	CATCAGATAG	TTCAACTTCA	CAGTCTAGCT	CAACCACTCC	1150
45	AAGCACAAAT	AATAGTACGA	CTACCAATCC	TAACAATAAT	ACGCAACAAT	1200
	CAAATACAAC	CCCTGATC				1218

## 50 2) INFORMATION FOR SEQ ID NO: 1008

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1223 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 60 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-05

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1008

```

5  GTAGACCAAG AAGCTCAAAA ACATCTGTGG GATATCTACA ACTCCGATCA      50
   ATACGTCTCT TACCCTGACG ATGATTTGCA AGTCGCATCT ACGGTCGTAG      100
   ATGTTTCAAA TGGTAAAGTC ATCGCCCAAC TTGGAGCTCG TCACCAAGCA      150
   AGTAACGTTT CATTTGGTAC CAACCAAGCT GTGGAAACCA ATCGTGACTG      200
10  GGGTTCTGCT ATGAAACCAA TCACCGATTA TGCACCTGCC ATAGAATACG      250
   GTGTTTATGA TTCCACTGCA ACTATGGTTA ATGATATTCC TTATAACTAT      300
   CCGGGAACAA GCACACCTGT CTACAACTGG GATAGAGCAT ATTTTCGGTAA      350
   TATTACTCTG CAATATGCTC TTCAACAATC ACGAAATGTC ACAGCCGTTG      400
   AGACTTTGAA TAAGGTCGGT CTAGATAGAG CTAAAACCTT CCTTAATGGT      450
15  CTTGGTATCG ACTATCCAAG CATGCATTAT GCAAACGCCA TTTCAAGTAA      500
   TACAACAGAA TCTAATAAAC AATACGGAGC AAGTAGTGAA AAAATGGCTG      550
   CTGCTTATGC TGCCTTTGCA AATGGTGGCA CTTACTATAA ACCAATGTAT      600
   ATCCATAAAG TCGTCTTCAG TGATGGAAGT AAAAAAGAGT TCTCTAATGT      650
   CGGAACTCGT GCCATGAAGG AAACGACAGC CTATATGATG ACCGACATGA      700
20  TGAAAACAGT CTTGACTTAT GGAAGTGGG CTTGGAGCCTA TCTTCCTTGG      750
   CTTCTCAAG CTGTTAAAAC AGGAACCTCT AACTATACAG ATGAGGAAGT      800
   TGAAAACCAAC ATCAAGAACA CTGGCTATGT AGCTCCAGAT GAAATGTTTG      850
   TTGGTTATAC TCGTAAGTAT TCTATGGCTG TATGGACAGG TTATTCGAAT      900
   CGTTTAACTC CTATCGTTGG AGATGGTTTC CTAGTTGCAG CTAAAGTTTA      950
25  TCGCTCAATG ATAACGTATC TATCAGAAGA TACTCATCCA GAAGACTGGA      1000
   CGATGCCAGA CGGACTTTTC AGAAACGGGG AATTTGTATT CAAAAATGGA      1050
   GCTCGCCCAA TATGGACTGA ACCCTCTACT CAACAATCCT CAACAGCTGA      1100
   AAGTTCAAGC TCATCATCAG ATAGTTCAAC TTCACAGTCT AGCTCAACCA      1150
   CTCCAAGCAC AAATAATAGT ACGACTACCA ATCCTAACAA TAATACGCAA      1200
30  CAATCAAATA CAACCCCTGA TCA                                     1223
  
```

## 2) INFORMATION FOR SEQ ID NO: 1009

```

35  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 1214 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
40  (D) TOPOLOGY: Linear

      (ii) MOLECULE TYPE: Genomic DNA

      (vi) ORIGINAL SOURCE:
45  (A) ORGANISM: Streptococcus pneumoniae
      (B) STRAIN: StrR-06
  
```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1009

```

50  ACCAAGAAGC TCAAAAACAT CTGTGGGATA TCTACAACTC CGATCAATAC      50
   GTCTCTTACC CTGACGATGA TTTGCAAGTC GCATCTACGG TCGTAGATGT      100
   TTCAAATGGT AAAGTCATCG CCCAACTTGG AGCTCGTCAC CAAGCAAGTA      150
   ACGTTTCATT TGGTACCAAC CAAGCTGTGG AAACCAATCG TGACTGGGGT      200
   TCTGCTATGA AACCAATCAC CGATTATGCA CCTGCCATAG AATACGGTGT      250
55  TTATGATTCC ACTGCAACTA TGGTTAATGA TATTCCTTAT AACTATCCGG      300
   GAACAAGCAC ACCTGTCTAC AACTGGGATA GAGCATATTT CGGTAATATT      350
   ACTCTGCAAT ATGCTCTTCA ACAATCACGA AATGTCACAG CCGTTGAGAC      400
   TTTGAATAAG GTCGGTCTAG ATAGAGCTAA AACCTTCCTT AATGGTCTTG      450
   GTATCGACTA TCCAAGCATG CATTATGCAA ACGCCATTTC AAGTAATACA      500
60  ACAGAATCTA ATAAACAATA CGGAGCAAGT AGTGAAAAAA TGGCTGCTGC      550
  
```

	TTATGCTGCC	TTTGCAAATG	GTGGCACTTA	CTATAAACCA	ATGTATATCC	600
	ATAAAGTCGT	CTTCAGTGAT	GGAAGTAAAA	AAGAGTTCTC	TAATGTCGGA	650
	ACTCGTGCCA	TGAAGGAAAC	GACAGCCTAT	ATGATGACCG	ACATGATGAA	700
	AACAGTCTTG	ACTTATGGAA	CTGGGCGTGG	AGCCTATCTT	CCTTGGCTTC	750
5	CTCAAGCTGG	TAAAACAGGA	ACCTCTAACT	ATACAGATGA	GGAAGTTGAA	800
	AACCACATCA	AGAACACTGG	CTATGTAGCT	CCAGATGAAA	TGTTTGTGG	850
	TTATACTCGT	AAGTATTCTA	TGGCTGTATG	GACAGGTTAT	TCGAATCGTT	900
	TAACTCCTAT	CGTTGGAGAT	GGTTTCCTAG	TTGCAGCTAA	AGTTTATCGC	950
	TCAATGATAA	CGTATCTATC	AGAAGATACT	CATCCAGAAG	ACTGGACGAT	1000
10	GCCAGACGGA	CTTTTCAGAA	ACGGGGAATT	TGTATTCAAA	AATGGAGCTC	1050
	GCCCAATATG	GACTGAACCC	TCTACTCAAC	AATCCTCAAC	AGCTGAAAGT	1100
	TCAAGCTCAT	CATCAGATAG	TTCAACTTCA	CAGTCTAGCT	CAACCACTCC	1150
	AAGCACAAAT	AATAGTACGA	CTACCAATCC	TAACAATAAT	ACGCAACAAT	1200
	CAAATACAAC	CCCT				1214

## 2) INFORMATION FOR SEQ ID NO: 1010

## 20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1223 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 25 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-07

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1010

	CCAAGAAGCT	CAAAAACATC	TGTGGGATAT	CTACAACTCC	GATCAATACG	50
35	TCTCTTACCC	TGACGATGAT	TTGCAAGTCG	CATCTACGGT	CGTAGATGTT	100
	TCAAATGGTA	AAGTCATCGC	ACAAC TTGGT	GCTCGTCATC	AAGCAAGTAA	150
	TGTTTCATTC	GGTACCAACC	AGGCCGTAGA	AACCAATCGT	GACTGGGGAT	200
	CATCAATGAA	ACCAATCACT	GACTATGCTC	CCGCTTTAGA	ATATGGAGTC	250
	TATGACTCTA	CTGCTTCTAT	TGTACATGAT	GTCCCTTATA	ACTATCCTGG	300
40	CACTGATACT	CCACTCTACA	ACTGGGATCA	TGTCTACTTT	GGAAACATTA	350
	CAATCCAGTA	TGCTCTTCAA	CAATCACGAA	ATGTCACAGC	CGTTGAGACT	400
	TTGAATAAGG	TCGGTCTAGA	TAGAGCTAAA	ACCTTCCTTA	ATGGTCTTGG	450
	TATCGACTAT	CCAAGCATGC	ATTATGCAAA	CGCCATTTCA	AGTAACACAA	500
	CTGAATCCAA	CAAAAAATAT	GGTGCAAGTA	GTGAAAAAAT	GGCTGCTGCC	550
45	TACGCTGCTT	TTGCTAATGG	TGGTATTTAT	CACAAACCAA	TGTATATCAA	600
	TAAATATCGTC	TTTAGTGATG	GTAGCGAAAA	AGAATTTTCT	GATGCTGGTA	650
	CACGAGCTAT	GAAAGAGACT	ACTGCCTATA	TGATGACTGA	AATGATGAAA	700
	ACTGTTTTAA	CTTACGGAAC	AGGACGTGGA	GCCTACCTAC	CATGGCTTCC	750
	ACAAGCAGGT	AAGACAGGTA	CTTCTAACTA	TACTGACGAA	GAAATTGAAA	800
50	AGTATATCAA	GAACACTGGT	TACGTAGCTC	CAGATGAAAT	GTTTGTAGGG	850
	TATACCCGTA	AATATGCAAT	GGCTGTTTGG	ACAGGATACT	CAAATCGTCT	900
	AACTCCAATC	ATCGGAGATG	GTTTCCTTGT	TGCTGGTAAA	GTCTATCGTT	950
	CAATGATAAC	TTACCTTTCT	GAAGATGACC	AACCTGGAGA	TTGGACAATG	1000
	CCAGATGGCT	TGTATAGAAA	TGGAGAATTC	GTATTTAAAA	ATGGTGCTCG	1050
55	TTCTACGTGG	AGCTCACCTG	CTCCACAACA	ACCCCCATCA	ACTGAAAGTT	1100
	CAAGCTCATC	ATCAGATAGT	TCAACTTCAC	AGTCTAACTC	AACCACTCCA	1150
	AGCACAAATA	ATAGTACGAC	TACCAATCCT	AACAATAATA	CGCAACAATC	1200
	AAATACAACC	CCTGATCAAC	AAA			1223

60

## 2) INFORMATION FOR SEQ ID NO: 1011

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1207 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-08

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1011

```

AAGCTCAAAA ACATCTGTGG GATATCTACA ACTCCGATCA ATACGTCTCT      50
TACCCTGACG ATGATTTGCA AGTCGCATCT ACGGTCGTAG ATGTTTCAAA      100
TGGTAAAGTC ATCGCCCAAC TTGGAGCTCG TCACCAAGCA AGTAACGTTT      150
CATTTGGTAC CAACCAAGCT GTGGAAACCA ATCGTGA CTG GGGTTCTGCT      200
ATGAAACCAA TCACCGATTA TGCACCTGCC ATAGAATACG GTGTTTATGA      250
TTCCACTGCA ACTATGGTTA ATGATATTCC TTATAACTAT CCGGGAACAA      300
GCACACCTGT CTACAACCTGG GATAGAGCAT ATTTTCGGTAA TATTACTCTG      350
CAATATGCTC TTCAACAATC ACGAAATGTC ACAGCCGTTG AGACTTTGAA      400
TAAGGTCGGT CTAGATAGAG CTAAAACCTT CCTTAATGGT CTTGGTATCG      450
ACTATCCAAG CATGCATTAT GCAAACGCCA TTTCAAGTAA TACAACAGAA      500
TCTAATAAAC AATACGGAGC AAGTAGTGAA AAAATGGCTG CTGCTTATGC      550
TGCCTTTGCA AATGGTGGCA CTTACTATAA ACCAATGTAT ATCCATAAAG      600
TCGTCTTCAG TGATGGAAGT AAAAAAGAGT TCTCTAATGT CGGAACTCGT      650
GCCATGAAGG AAACGACAGC CTATATGATG ACCGACATGA TGAAAACAGT      700
CTTGACTTAT GGAAGTGGG GTGGAGCCTA TCTTCCTTGG CTTCTCAAG      750
CTGGTAAAC AGGAACCTCT AACTATACAG ATGAGGAAGT TGAAAACCAC      800
ATCAAGAACA CTGGCTATGT AGCTCCAGAT GAAATGTTTG TTGGTTATAC      850
TCGTAAGTAT TCTATGGCTG TATGGACAGG TTATTCGAAT CGTTTAACTC      900
CTATCGTTGG AGATGGTTTC CTAGTTGCAG CTAAAGTTTA TCGCTCAATG      950
ATAACGTATC TATCAGAAGA TACTCATCCA GAAGACTGGA CGATGCCAGA     1000
CGGACTTTTC AGAAACGGGG AATTTGTATT CAAAATGGA GCTCGCCCAA     1050
TATGGACTGA ACCCTCTACT CAACAATCCT CAACAGCTGA AAGTTCAAGC     1100
TCATCATCAG ATAGTTCAAC TTCACAGTCT AGCTCAACCA CTCCAAGCAC     1150
AAATAATAGT ACGACTACCA ATCCTAACAA TAATACGCAA CAATCAAATA     1200
CAACCCC                                           1207

```

## 2) INFORMATION FOR SEQ ID NO: 1012

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1201 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-09

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1012

```

5  AAGCTCAAAA ACATCTGTGG GATATCTACA ACTCCGATCA ATACGTCTCT 50
   TACCCTGACG ATGATTTGCA AGTCGCATCT ACGGTCGTAG ATGTTTCAAA 100
   TGGTAAAGTC ATCGCCCAAC TTGGAGCTCG TCACCAAGCA AGTAACGTTT 150
   CATTTGGTAC CAACCAAGCT GTGGAAACCA ATCGTGACTG GGGTCTGTCT 200
   ATGAAACCAA TCACCGATTA TGCACCTGCC ATAGAATACG GTGTTTATGA 250
   TTCCACTGCA ACTATGGTTA ATGATATTCCT TTATAACTAT CCGGGAACAA 300
   GCACACCTGT CTACAACTGG GATAGAGCAT ATTTCCGTAA TATTACTCTG 350
   CAATATGCTC TTCAACAATC ACGAAATGTC ACAGCCGTTG AGACTTTGAA 400
10 TAAGGTCGGT CTAGATAGAG CTAAAACCTT CCTTAATGGT CTTGGTATCG 450
   ACTATCCAAG CATGCATTAT GCAAACGCCA TTTCAAGTAA TACAACAGAA 500
   TCTAATAAAC AATACGGAGC AAGTAGTGAA AAAATGGCTG CTGCTTATGC 550
   TGCCTTTGCC AACGGTGGAA TTTACCACAA ACCCATGTAT ATCAATAAGG 600
   TCGTCTTCAG TGACGGTAGT AAAAAAGAAT TTTCAGATGT AGGTACACGA 650
15 GCTATGAAAG AAACAACCTG TTACATGATG ACCGAAATGA TGAAAACCTGT 700
   CTTGGCATAAC GGAACCTGGT GTGGAGCCTA TCTCCCATGG TTAGCGCAAG 750
   CTGGTAAGAC AGGTACTTCT AACTACACAG ATGATGAAAT TGAAAAACAC 800
   ATCAAGAACA CTGGCTATGT AGCTCCAGAT GAAATGTTTG TTGGTTATAC 850
   TCGTAAGTAT TCTATGGCTG TATGGACAGG TTATTGCAAT CGTTTAACTC 900
20 CTATCGTTGG AGATGGTTTC CTAGTTGCAG CTAAAGTTTA TCGCTCAATG 950
   ATAACGTATC TATCAGAAGA TACTCATCCA GAAGACTGGA CGATGCCAGA 1000
   CGGACTTTTC AGAAACGGGG AATTTGTATT CAAAAATGGA GCTCGTTCTA 1050
   CGTGGAACCT ACCTGCTCCA CAACAACCCC CATCAACTGA AAGTTCAAGC 1100
   TCATCATCAG ATAGTTCAAC TTCACAGTCT AGCTCAACCA CTCCAAGCAC 1150
25 AAATAATAGT ACGACTACCG ATCCTAACAA TAATACGCAA CAATCAAATA 1200
   C 1201

```

30 2) INFORMATION FOR SEQ ID NO: 1013

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 1220 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

40 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1013

```

45 TGTAGACCAA GAAGCTCAAA AACATCTGTG GGATATCTAC AACTCCGATC 50
   AATACGTCTC TTACCCTGAC GATGATTTGC AAGTCGCATC TACGGTCGTA 100
   GATGTTTCAA ATGGTAAAGT CATCGCACAA CTTGGTGCTC GTCATCAAGC 150
   AAGTAATGTT TCATTCCGTA CCAACCAGGC CGTAGAAACC AATCGTGACT 200
50 GGGGATCATC AATGAAACCA ATCACTGACT ATGCTCCCGC TTTAGAATAT 250
   GGAGTCTATG ACTCTACTGC TTCTATTGTA CATGATGTCC CTTATAACTA 300
   TCCTGGCACT GATACTCCAC TCTACAACCT GGATCATGTC TACTTTGGAA 350
   ACATTACAAT CCAGTATGCT CTTCAACAAT CACGAAATGT CACAGCCGTT 400
   GAGACTTTGA ATAAGGTCGG TCTAGATAGA GCTAAAACCT TCCTTAATGG 450
55 TCTTGGTATC GACTATCCAA GCATGCATTA TGCAAACGCC ATTTCAAGTA 500
   ACACAACCTGA ATCCAACAAA AAATATGGTG CAAGTAGTGA AAAAATGGCT 550
   GCTGCCTACG CTGCTTTTGC TAATGGTGGT ATTTATCACA AACCAATGTA 600
   TATCAATAAA ATCGTCTTTA GTGATGGTAG CGAAAAAGAA TTTTCTGATG 650
   CTGGTACACG AGCTATGAAA GAGACTACTG CCTATATGAT GACTGAAATG 700
60 ATGAAAACCTG TTTTAACTTA CGGAACAGGA CGTGGAGCCT ACCTACCATG 750

```

	GCTTCCACAA	GCAGGTAAGA	CAGGTACTTC	TAAGTATACT	GACGAAGAAA	800
	TTGAAAAGTA	TATCAAGAAC	ACTGGCTACG	TAGCTCCAGA	TGAAATGTTT	850
	GTGGGTATA	CTCGTAAGTA	TTCTATGGCT	GTATGGACAG	GTTATTTCGAA	900
	TCGTTTAACT	CCTATCGTTG	GAGATGGTTT	CCTAGTTGCA	GCTAAAGTTT	950
5	ATCGCTCTAT	GATGACCTAC	CTGTCTGAAG	GAAGCAATCC	AGAGGATTGG	1000
	AATATACCAG	AGGGGCTCTA	CAGAAATGGA	GAATTCGTAT	TAAAAATGG	1050
	TGCTCGTTCT	ACGTGGAGCT	CACCTGCTCC	ACAACAACCC	CCATCAACTG	1100
	AAAGTTCAAG	CTCATCATCA	GATAGTTCAA	CTTCACAGTC	TAGCTCAACC	1150
	ACTCCAAGCA	CAAATAATAG	TACGACTACC	AATCCTAACA	ATAATACGCA	1200
10	ACAATCAAAT	ACAACCCCTG				1220

## 2) INFORMATION FOR SEQ ID NO: 1014

15

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1199 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

20

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-11

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1014

30	CAAAAACATC	TGTGGGATAT	TTACAATACA	GACGAATACG	TTGCCTATCC	50
	AGACGATGAA	TTGCAAGTCG	CTTCTACCAT	TGTTGATGTT	TCTAACGGTA	100
	AAGTCATTGC	CCAGCTAGGA	GCACGCCATC	AGTCAAGTAA	TGTTTCCTTC	150
	GGAATTAACC	AAGCAGTAGA	AACAAACCGC	GACTGGGGAT	CAACTATGAA	200
	ACCGATCACA	GACTATGCTC	CTGCCTTGGA	GTACGGTGTC	TACGATTCAA	250
35	CTGCTACTAT	CGTTCACGAT	GAGCCCTATA	ACTACCCTGG	GACAAATACC	300
	CCTGTTTATA	ACTGGGATAG	GGGCTACTTT	GGCAACATCA	CCTTGCAATA	350
	CGCCCTGCAA	CAATCGCGAA	ACGTCCCAGC	CGTGGAAACT	CTAAACAAGG	400
	TCGGACTCAA	CCGCGCCAAG	ACTTTCCTAA	ATGGTCTCGG	AATCGACTAC	450
	CCAAGTATTC	ACTACTCAA	TGCCATTTCA	AGTAACACAA	CCGAATCAGA	500
40	CAAAAAATAT	GGAGCAAGTA	GTGAAAAGAT	GGCTGCTGCT	TACGCTGCCT	550
	TTGCAAATGG	TGGAACCTAC	TATAAACCAA	TGTATATCCA	TAAAGTCGTC	600
	TTTAGTGATG	GGAGTGAAAA	AGAGTTCTCT	AATGTCGGAA	CTCGTGCCAT	650
	GAAGGAAACG	ACAGCCTATA	TGATGACCGA	CATGATGAAA	ACTGTCCTAG	700
	TATACGGAAT	CGGACGTGGA	GCCTACCTAC	CTTGGCTTCC	ACAAGCAGGT	750
45	AAAACAGGTA	CTTCTAACTA	TACTGACGAA	GAAATTGAAA	AGTATATCAA	800
	GAACACTGGT	TACGTAGCTC	CAGATGAAAT	GTTTGTAGGG	TATACCCGCA	850
	AATATGCAAT	GGCTGTATGG	ACAGGCTATT	CTAACCGTCT	GACACCACTT	900
	GTAGGCGATG	GCCTTACGGT	CGCTGCTAAA	GTTTACCGCT	CTATGATGAC	950
	CTACCTGTCT	GAAGGAAGCA	ATCCAGAGGA	TTGGAATATA	CCAGAGGGGC	1000
50	TCTACAGAAA	TGGAGAATTC	GTATTTAAAA	ATGGTGCTCG	TTCTACGTGG	1050
	AACTCACCTG	CTCCACAACA	ACCCCATCA	ACTGAAAGTT	CAAGCTCATC	1100
	ATCAGATAGT	TCAACTTCAC	AGTCTAGCTC	AACCACTCCA	AGCACAAATA	1150
	ATAGTACGAC	TACCAATCCT	AACAATAATA	CGCAACAATC	AAATACAAC	1199

55

## 2) INFORMATION FOR SEQ ID NO: 1015

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1211 bases

60

(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-12

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1015

	AAGCTCAAAA	ACATCTGTGG	GATATCTACA	ACTCCGATCA	ATACGTCTCT	50
	TACCCTGACG	ATGATTTGCA	AGTCGCATCT	ACGGTCGTAG	ATGTTTCAAA	100
15	TGGTAAAGTC	ATCGCCCAAC	TTGGAGCTCG	TCACCAAGCA	AGTAACGTTT	150
	CATTTGGTAC	CAACCAAGCT	GTGGAAACCA	ATCGTGACTG	GGGTTCTGCT	200
	ATGAAACCAA	TCACCGATTA	TGCACCTGCC	ATAGAATACG	GTGTTTATGA	250
	TTCCACTGCA	ACTATGGTTA	ATGATATTCC	TTATAACTAT	CCGGGAACAA	300
	GCACACCTGT	CTACCACTGG	GATAGAGCAT	ATTTCCGGTAA	TATTACTCTG	350
20	CAATATGCTC	TTCAACAATC	ACGAAATGTC	ACAGCCGTTG	AGACTTTGAA	400
	TAAGGTCGGT	CTAGATAGAG	CTAAAACCTT	CCTTAATGGT	CTTGGTATCG	450
	ACTATCCAAG	CATGCATTAT	GCAAACGCCA	TTTCAAGTAA	TACAACAGAA	500
	TCTAATAAAC	AATACGGAGC	AAGTAGTGAA	AAAATGGCTG	CTGCTTATGC	550
	TGCCTTTGCC	AACGGTGGAA	TTTACCACAA	ACCCATGTAT	ATCAATAAGG	600
25	TCGTCTTCAG	TGACGGTAGT	AAAAAAGAAT	TTTCAGATGT	AGGTACACGA	650
	GCTATGAAAG	AAACAACCTG	TTACATGATG	ACCGAAATGA	TGAAAACCTGT	700
	CTTGGCATA	GGAACCTGGC	GTGGAGCCTA	TCTCCCATGG	TTAGCGCAAG	750
	CTGGTAAGAC	AGGTACTTCT	AACTACACAG	ATGATGAAAT	TGAAAAACAC	800
	ATCAAGAACA	CTGGCTATGT	AGCTCCAGAT	GAAATGTTTG	TTGGTTATAC	850
30	TCGTAAGTAT	TCTATGGCTG	TATGGACAGG	TTATTCGAAT	CGTTTAACTC	900
	CTATCGTTGG	AGATGGTTTC	CTAGTTGCAG	CTAAAGTTTA	TCGCTCAATG	950
	ATAACGTATC	TGTCTGAAGG	AAGCAATCCA	GAGGATTGGA	ATATACCAGA	1000
	GGGGCTCTAC	AGAAATGGAG	AATTCGTATT	TAAAAATGGT	GCTCGTTCTA	1050
	CGTGGAGCTC	ACCTGCTCCA	CAACAACCCC	CATCAACTGA	AAGTTCAAGC	1100
35	TCATCATCAG	ATAGTTCAAC	TTCACAGTCT	AACTCAACCA	CTCCAAGCAC	1150
	AAATAATAGT	ACGACTACCA	ATCCTAACAA	TAATACGCAA	CAATCAAATA	1200
	CAACCCCTGA	T				1211

40

2) INFORMATION FOR SEQ ID NO: 1016

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1222 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

50

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-13

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1016

	GTAGACCAAG	AAGCTCAAAA	ACATCTGTGG	GATATCTACA	ACTCCGATCA	50
	ATACGTCTCT	TACCCTGACG	ATGATTTGCA	AGTCGCATCT	ACGGTCGTAG	100
	ATGTTTCAAA	TGGTAAAGTC	ATCGCACAAAC	TTGGTGCTCG	TCATCAAGCA	150
60	AGTAATGTTT	CATTCGGTAC	CAACCAGGCC	GTAAGAAACCA	ATCGTGACTG	200

```

GGGATCATCA ATGAAACCAA TCACTGACTA TGCTCCCGCT TTAGAATATG 250
GAGTCTATGA CTCTACTGCT TCTATTGTAC ATGATGTCCC TTATAACTAT 300
CCTGGCACTG ATACTCCACT CTACAACCTG GATCATGTCT ACTTTGGAAA 350
CATTACAATC CAGTATGCTC TTCAACAATC ACGAAATGTC ACAGCCGTTG 400
5 AGACTTTGAA TAAGGTCGGT CTAGATAGAG CTAAAACCTT CCTTAATGGT 450
CTTGGTATCG ACTATCCAAG CATGCATTAT GCAAACGCCA TTTCAAGTAA 500
CACAACCTGAA TCCAACAAAA AATATGGTGC AAGTAGTGAA AAAATGGCTG 550
CTGCCTACGC TGCTTTTGCT AATGGTGGTA TTTATCACAA ACCAATGTAT 600
ATCAATAAAA TCGTCTTTAG TGATGGTAGC GAAAAAGAAT TTTCTGATGC 650
10 TGGTACACGA GCTATGAAAG AGACTACTGC CTATATGATG ACTGAAATGA 700
TGAAAACCTGT TTAACTTAC GGAACAGGAC GTGGAGCCTA CCTACCATGG 750
CTTCCACAAG CAGGTAAGAC AGGTACTTCT AACTATACTG ACGAAGAAAT 800
TGAAAAGTAT ATCAAGAACA CTGGCTACGT AGCTCCAGAT GAAATGTTTG 850
TGGGTTATAC TCGTAAGTAT TCTATGGCTG TATGGACAGG TTATTGCAAT 900
15 CGTTTAACTC CTATCGTTGG AGATGGTTTC CTAGTTGCAG CTAAAGTTTA 950
TCGCTCTATG ATGACCTACC TGTCTGAAGG AAGCAATCCA GAGGATTGGA 1000
ATATACCAGA GGGGCTCTAC AGAAATGGAG AATTCGTATT TAAAAATGGT 1050
GCTCGTTCTA CGTGGAGCTC ACCTGCTCCA CAACAACCCC CATCAACTGA 1100
AAGTTCAAGC TCATCATCAG ATAGTTCAAC TTCACAGTCT AGCTCAACCA 1150
20 CTCCAAGCAC AAATAATAGT ACGACTACCA ATCCTAACAA TAATACGCAA 1200
CAATCAAATA CAACCCCTGA TC 1222

```

## 25 2) INFORMATION FOR SEQ ID NO: 1017

## (i) SEQUENCE CHARACTERISTICS:

```

30 (A) LENGTH: 1229 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

## 35 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Streptococcus pneumoniae
    (B) STRAIN: StrR-14

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1017

```

40 TGTAGACCAA GAAGCTCAAA AACATCTGTG GGATATTTAC AATACAGACG 50
AATACGTTGC CTATCCAGAC GATGAATTGC AAGTCGCTTC TACCATTGTT 100
GATGTTTCTA ACGGTAAAGT CATTGCCACG CTAGGAGCAC GCCATCAGTC 150
AAGTAATGTT TCCTTCGGAA TTAACCAAGC AGTAGAAACA AACCGCGACT 200
45 GGGGATCAAC TATGAAACCG ATCACAGACT ATGCTCCTGC CTTGGAGTAC 250
GGTGTCTACG ATTCAACTGC TACTATCGTT CACGATGAGC CCTATAACTA 300
CCCTGGGACA AATACCCCTG TTTATAACTG GGATAGGGGC TACTTTGGCA 350
ACATCACCTT GCAATACGCC CTGCAACAAT CGCGAAACGT CCCAGCCGTG 400
GAAACTCTAA ACAAGGTCGG ACTCAACCGC GCCAAGACTT TCCTAAATGG 450
50 TCTCGGAATC GACTACCCAA GTATTCACTA CTCAAATGCC ATTTCAAGTA 500
ACACAACCGA ATCAGACAAA AAATATGGAG CAAGTAGTGA AAAGATGGCT 550
GCTGCTTACG CTGCCTTTGC AAATGGTGGA ACTTACTATA AACCAATGTA 600
TATCCATAAA GTCGTCTTTA GTGATGGGAG TGAAAAAGAG TTCTCTAATG 650
TCGGAACCTG TGCCATGAAA GAAACAACCTG CTTACATGAT GACCGAAATG 700
55 ATGAAAACCTG TCCTGGCATA CGGAAGTGGT CGTGGAGCCT ATCTCCCATG 750
GTTAGCGCAA GCTGGTAAGA CAGGTACTTC TAACTACACA GATGATGAAA 800
TTGAAAAACA CATCAAGAAC ACTGGCTATG TAGCTCCAGA TGAAATGTTT 850
TTGGTTTATA CCTGTAAGTA TTCTATGGCT TATGGACAG GTTATTCGAA 900
TCTTTTAACT CCTATCGTTG GAGATGGTTT CCTAGTTGCA GCTAAAGTTT 950
60 ATCGCTCAAT GATAACGTAT CTATCAAAAG ATACTCATCC AGAAGACTGG 1000

```



	ACGATGCCAG	ACGGACTTTT	CAGAAACGGG	GAATTTGTAT	TCAAAAATGG	1050
	AGCTCGTTCT	ACGTGGAAC	CACCTGCTCC	ACAACAACCC	CCATCAACTG	1100
	AAAGTTCAAG	CTCATCATCA	GATAGTTCAA	CTTCACAGTC	TAACTCAACC	1150
	ACTCCAAGCA	CAAATAATAG	TACGACTACC	AATCCTAACA	ATAATACGCA	1200
5	ACAATCAAAT	ACAACCCCTG	ATCAACAAA			1229

## 2) INFORMATION FOR SEQ ID NO: 1018

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1225 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1018

25	ATGTAGACCA	AGAGGCTCAA	AAACGTCTGT	GGGATATCTA	CAACTCCGAT	50
	CAATACGTCT	CTTACCCTGA	CGATGATTTG	CAAGTCGCAT	CTACGGTCGT	100
	AGATGTTTCA	AATGGTAAAG	TCATCGCCCA	ACTTGGAGCT	CGTCACCAAG	150
	CAAGTAACGT	TTCATTTGGT	ACCAACCAAG	CTGTGGAAAC	CAACCGTGAC	200
	TGGGGATCAA	GCATGAAACC	AATCACTGAT	TATGCCCCAG	CCTTAGAATA	250
30	TGGTGTATAT	GATTCCACTG	CAACTATGGT	TAATGATATT	CCTTATAACT	300
	ATCCGGGCAC	AAGCACACCT	GTCTACAAC	GGGATCGAGC	ATATTTTGGT	350
	AATATTAGCC	TGCAATATGC	CCTTCAACAA	TCTCGTAACG	TGCTTGCCGT	400
	TGAAACACTA	AACAAGGTTG	GTTTAGATAG	AGCCAAAAC	TTCCTAAATG	450
	GTTTGGGAAT	TGACTATCCA	AGTATTCACT	ACTCAAATGC	TATTTCAAGT	500
35	AATACAAC	AATCTAGTAA	ACAGTACGGG	GCAAGCAGTG	AGAAAATGGC	550
	TGCGGCTTAC	GCTGCATTCT	CTAATGGCGG	TATTTACCAC	AAACCAATGT	600
	ACATCAATAA	AGTTGTCTTT	AGTGATGGTA	GCGAAAAAGA	ATTTTCTGAT	650
	GCTGGTACAC	GAGCTATGAA	AGAGACTACT	GCCTATATGA	TGACTGAAAT	700
	GATGAAAAC	GTTTTAACTT	ACGGAACAGG	ACGTGGAGCC	TACCTACCAT	750
40	GGCTTCCACA	AGCAGGTAAG	ACAGGTACTT	CTAACTATAC	TGACGACGAA	800
	ATTGAAAAGT	ATATCAAGAA	CACTGGCTAC	GTAGCTCCAG	ATGAAATGTT	850
	TGTGGGTTAT	ACTCGTAAAG	ATTCTATGGC	TGTATGGACT	GGATACTCAA	900
	ATCGTTTAAC	TCCAATCATT	GGAGATGGTT	TCCTAGTTGC	TGCCAAAGTT	950
	TATCGCTCAA	TGATATCGTA	TCTATCAGAA	GATGACCATC	CTGGAGATTG	1000
45	GACAATGCCT	GAGGGAGTAT	ACAGAAAGTG	AGAATTTCGT	TTTAAAAATG	1050
	GTGCTCGTTC	TACGTGGAGC	TCCCCTGCTC	CACAACAACC	CCCATCAACT	1100
	GAAAGTTCAA	GCTCATCATC	AGATAGTTCA	ACTTCACAGT	CTAGCTTAAC	1150
	CACTCCAAGC	ACAAATAATA	GTACGACTAC	CAATCCTAAC	AACAATACGC	1200
	AACAATCAAA	TACAACCCCT	GATCA			1225

## 2) INFORMATION FOR SEQ ID NO: 1019

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1439 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: StrR-01

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1019

```

10 GCCTCTATTT CAAAGGAGAT GCCTGGCATT AGTATTTCTA CTTCTTGGA 50
   TCGAAAGGTT TTGGAAACTT CCCTTTCTTC TATAGTAGGG AGTGATATCCA 100
   GTGAAAAAGC TGGTCTCCCA GCGGAAGAAG CAGAATCCTA TCTTAAAAAA 150
   GGCTATTCTC TAAATGACCG TGTTGGAACC TCCTATTTGG AAAAGCAATA 200
   TGAAGAGACC TTACAAGGAA AACGCTCGGT AAAAGAAATC CATCTGGATA 250
   AATATGGCAA TATGGAAAGC GTGGACACAA TTGAGGAAGG TAGTAAGGGA 300
15 AACAAATATCA AACTGACCAT TGATTTGGCC TTCCAAGATA GCGTGGATGC 350
   TTTGCTGAAA AGTTATTTCA ATTCCGAGCT AGGAAATGGT GGAGCTAAAT 400
   ATTCTGAAGG TGTCTATGCA GTCGCCCTTA ACCCAAAAAC AGGTGCTGTT 450
   TTATCCATGT CAGGGATCAA ACATGACCTG AAAACGGGAG AGTTGACTCC 500
   TGATTCCTTG GGAACGGTAA CCAATGTCTT TGTCCCAGGT TCGGTTGTCA 550
20 AGGCTGCGAC CATCAGCTCA GGTGGGAAA ATGGTGT TTT ATCAGGAAAC 600
   CAAACCTTAA CAGATCAGCC TATTGTTTTT CAAGGTTTCA CTCCAATTTA 650
   TTCTTGGTAT AAATTGGCAT ATGGATCTTT TCCTATTACA GCTGTGGAAG 700
   CCTTGAGTA TTCATCCAAT GCTTACATGG TTCAAACCGC TCTTGAATC 750
   ATGGGCCAGA CCTATCAACC AAATATGTTT GTTGAACCA GCAATTGGA 800
25 AACAGCTATG GGAAACTTC GTGCGACCTT TGGCGAATAT GGCTTGGGGG 850
   CTGCGACCGG AATTGACCTA CCAGATGAAT CTAAGGATT TGTTCCTCAA 900
   GAGTATAGCT TTGCTAATTA CATCACC AAT TCCTTTGGGC AGTTTGATAA 950
   CTATACGCC ATGCAGTTGG CTCAGTATGT AGCAACTATT GCAAATAATG 1000
   GTGTTTCGTG GGTCTCTCGT ATTGTTGAAG GCATTTATGG TAATAATGAT 1050
30 AAGGGAGGAC TGGGTGACTT GATTCAGCAA CTGCAACCGA CAGAGATGAA 1100
   TAAGGTCAAT ATATCCGACT CCGATATGAG CATCTTGCAC CAAGGTTTTT 1150
   ATCAGGTTGC CCATGGTACT AGTGGATTGA CAACTGGACG TGCCTTTTCA 1200
   AATGGTGCCT TGGTATCCAT TAGCGGAAAA ACAGGTACAG CCGAAAGCTA 1250
   TGTGGCAGAT GGTCAGCAAG CAACCAATAC CAATGCGGTG GCCTATGCCC 1300
35 CATCTGATAA TCCCCAAATC GCTGTGCGAC TGGTCTTTCC TCATAATACC 1350
   AATCTAACAA ATGGTGATG ACCTTCCATT GCGCGTGACA TTATCAATCT 1400
   GTATCAAAAA TACCATCCAA TGAAC TAGAA AGGAAATTA 1439

```

40

2) INFORMATION FOR SEQ ID NO: 1020

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1441 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: StrR-02

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1020

```

60 GCCTCTATTT CAAAGGAGAT GCCTGGCATT AGTATTTCTA CTTCTTGGA 50
   TCGAAAGGTT TTGGAAACTT CCCTTTCTTC TATAGTAGGG AGTGATATCCA 100
   GTGAAAAAGC TGGTCTCCCA GCGGAAGAAG CAGAATCCTA TCTTAAAAAA 150
   GGCTATTCTC TAAATGACCG TGTTGGAACC TCCTATTTGG AAAAGCAATA 200

```

	TGAAGAGACC	TTACAAGGAA	AACGCTCGGT	AAAAGAAATC	CATCTGGATA	250
	AATATGGCAA	TATGGAAAGC	GTGGACACAA	TTGAGGAAGG	TAGTAAGGGA	300
	AACAATATCA	AACTGACCAT	TGATTTGGCC	TTCCAAGATA	GCGTGGATGC	350
	TTTGTCTGAAA	AGTTATTTCA	ATTCCGAGCT	AGGAAATGGT	GGAGCTAAAT	400
5	ATTCTGAAGG	TGCTATGCA	GTCGCCCTTA	ACCCAAAAAC	AGGTGCTGTT	450
	TTATCCATGT	CAGGGATCAA	ACATGACCTG	AAAACGGGAG	AGTTGACTCC	500
	TGATTCCTTG	GGAACGGTAA	CCAATGTCTT	TGTCCCAGGT	TCGGTTGTCA	550
	AGGCTGCGAC	CATCAGCTCA	GGTTGGGAAA	ATGGTGTTTT	ATCAGGAAAC	600
	CAAACCTTAA	CAGATCAGCC	TATTGTTTTT	CAAGGTTTCA	CTCCAATTTA	650
10	TTCTTGGTAT	AAATTGGCAT	ATGGATCTTT	TCCTATTACA	GCTGTGGAAG	700
	CCTTGGAGTA	TTCATCCAAT	GCTTACATGG	TTCAAACCGC	TCTTGGAATC	750
	ATGGGCCAGA	CCTATCAACC	AAATATGTTT	GTTGGAACCA	GCAATTTGGA	800
	AACAGCTATG	GGAAAACTTC	GTGCGACCTT	TGGCGAATAT	GGCTTGGGGG	850
	CTGCGACCGG	AATTGACCTA	CCAGATGAAT	CTACTGGATT	TGTTCCCAAA	900
15	GAGTATAGCT	TTGCTAATTA	CATCACCAAT	TCCTTTGGGC	AGTTTGATAA	950
	CTATACGCCC	ATGCAGTTGG	CTCAGTATGT	AGCAACTATT	GCAAATAATG	1000
	GTGTTCTGAT	GGCTCCTCGT	ATTGTTGAAG	GCATTTATGG	TAATAATGAT	1050
	AAGGGAGGAC	TGGGTGACTT	GATTTCAGCA	CTGCAACCGA	CAGAGATGAA	1100
	TAAGGTCAAT	ATATCCGACT	CCGATATGAG	CATCTTGCAC	CAAGGTTTTT	1150
20	ATCAGGTTGC	CCATGGTACT	AGTGGATTGA	CAACTGGACG	TGCCTTTTCA	1200
	AATGGTGCC	TGGTATCCAT	TAGCGGAAAA	ACAGGTACAG	CCGAAAGCTA	1250
	TGTGGCAGAT	GGTCAGCAAG	CAACCAATAC	CAATGCGGTG	GCCTATGCCC	1300
	CATCTGATAA	TCCCCAAATC	GCTGTCGCAG	TGGTCTTTCC	TCATAATACC	1350
	AATCTAACAA	ATGGTGTTAG	ACCTTCCATT	GCGCGTGACA	TTATCAATCT	1400
25	GATCAAAAA	TACCATCCAA	TGAAC TAGAA	AGGAAATTAT	G	1441

## 2) INFORMATION FOR SEQ ID NO: 1021

30

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1396 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

35

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-03

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1021

45	TATTGCCTCT	ATTTCAAAGG	AGATGCCTGG	CATTAGTATT	TCTACTTCTT	50
	GGGATCGAAA	GGTTTTGGAA	ACTTCCCTTT	CTTCTATAGT	AGGGAGTGTA	100
	TCCAGTGAAA	AAGCTGGTCT	CCCAGCGGAA	GAAGCAGAAT	CCTATCTTAA	150
	AAAAGGCTAT	TCTCTAAATG	ACCGTGTTGG	AACCTCCTAT	TTGGAAAAGC	200
	AATATGAAGA	GACCTTACAA	GGAAAACGCT	CGGTAAAAGA	AATCCATCTG	250
50	GATAAATATG	GCAATATGGA	AAGCGTGGAC	ACAATTGAGG	AAGGTAGTAA	300
	GGGAAACAAT	ATCAAACTGA	CCATTGATTT	GGCCTTCCAA	GATAGCGTGG	350
	ATGCTTTGCT	GAAAAGTTAT	TTCAATTCCG	AGCTAGGAAA	TGGTGGAGCT	400
	AAATATTCTG	AAGGTGTCTA	TGCAGTCGCC	CTTAACCCAA	AAACAGGTGC	450
	TGTTTTATCC	ATGTCAGGGA	TCAAACATGA	CCTGAAAACG	GGAGAGTTGA	500
55	CTCCTGATTC	CTTGGGAACG	GTAACCAATG	TCTTTGTCCC	AGGTTTCGGTT	550
	GTCAAGGCTG	CGACCATCAG	CTCAGGTTGG	GAAAATGGTG	TTTTATCAGG	600
	AAACCAAACC	TTAACAGATC	AGCCTATTGT	TTTCCAAGGT	TCAGCTCCAA	650
	TTTATTCTTG	GTATAAATTG	GCATATGGAT	CTTTTCCTAT	TACAGCTGTG	700
	GAAGCCTTGG	AGTATTCATC	CAATGCCTTAC	ATGGTTCAAA	CCGCTCTTGG	750
60	AATCATGGGC	CAGACCTATC	AACCAAATAT	GTTTGTGGA	ACCAGCAATT	800

	TGGAACACAGC	TATGGGAAAA	CTTCGTGCGA	CCTTTGGCGA	ATATGGCTTG	850
	GGGGCTGCGA	CCGGAATTGA	CCTACCAGAT	GAATCTACTG	GATTGTGTTCC	900
	CAAAGAGTAT	AGCTTTGCTA	ATTACATCAC	CAATTCCTTT	GGGCAGTTTG	950
	ATAACTATAC	GCCCCATGCAG	TTGGCTCAGT	ATGTAGCAAC	TATTGCAAAT	1000
5	AATGGTGTTT	GTGTGGCTCC	TCGTATTGTT	GAAGGCATTT	ATGGTAATAA	1050
	TGATAAGGGA	GGACTGGGTG	ACTTGATTCA	GCAACTGCAA	CCGACAGAGA	1100
	TGAATAAGGT	CAATATATCC	GACTCCGATA	TGAGCATCTT	GCACCAAGGT	1150
	TTTTATCAGG	TTGCCCATGG	TACTAGTGGA	TTGACAACCTG	GACGTGCCTT	1200
	TTCAAATGGT	GCCTTGGTAT	CCATTAGCGG	AAAAACAGGT	ACAGCCGAAA	1250
10	GCTATGTGGC	AGATGGTCAG	CAAGCAACCA	ATACCAATGC	GGTGGCCTAT	1300
	GCCCCATCTG	ATAATCCCCA	AATCGCTGTC	GCAGTGGTCT	TTCTCATATA	1350
	TACCAATCTA	ACAAATGGTG	TAGGACCTTC	CATTGCGCGT	GACATT	1396

15

## 2) INFORMATION FOR SEQ ID NO: 1022

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1428 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-04

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1022

	AAAGGAGATG	CCTGGCATTG	GTATTTCTAC	TTCTTGGGAT	CGAAAGGTTT	50
	TGGAAACTTC	CCTTTCTTCT	ATAGTAGGGA	GTGTATCCAG	TGAAAAAGCT	100
	GGTCTCCCAG	CGGAAGAAGC	AGAATCCTAT	CTTAAAAAAG	GCTATTCTCT	150
35	AAATGACCGT	GTTGGAACCT	CCTATTGGGA	AAAGCAATAT	GAAGAGACCT	200
	TACAAGGAAA	ACGCTCGGTA	AAAGAAATCC	ATCTGGATAA	ATATGGCAAT	250
	ATGGAAAGCG	TGGACACAAT	TGAGGAAGGT	AGTAAGGGAA	ACAATATCAA	300
	ACTGACCATT	GATTTGGCCT	TCCAAGATAG	CGTGGATGCT	TTGCTGAAAA	350
	GTTATTTCAA	TTCCGAGCTA	GGAAATGGTG	GAGCTAAATA	TTCTGAAGGT	400
40	GTCTATGCAG	TCGCCCTTAA	CCCAAAAACA	GGTGCTGTTT	TATCCATGTC	450
	AGGGATCAAA	CATGACCTGA	AAACGGGAGA	GTTGACTCCT	GATTCCCTGG	500
	GAACGGTAAC	CAATGTCTTT	GTCCCAGGTT	CGGTTGTCAA	GGCTGCGACC	550
	ATCAGCTCAG	GTTGGGAAAA	TGGTGTTTTA	TCAGGAAACC	AAACCTTAAC	600
	AGATCAGCCT	ATTGTTTTTC	AAGGTTTCAGC	TCCAATTTAT	TCTTGGTATA	650
45	AATTGGCATA	TGGATCTTTT	CCTATTACAG	CTGTGGAAGC	CTTGGAGTAT	700
	TCATCCAATG	CTTACATGGT	TCAAACCGCT	CTTGGAAATCA	TGGGCCAGAC	750
	CTATCAACCA	AATATGTTTG	TTGGAACCAAG	CAATTTGGAA	ACAGCTATGG	800
	GAAAACCTCG	TGCGACCTTT	GGCGAATATG	GCTTGGGGGC	TGCGACCGGA	850
	ATTGACCTAC	CAGATGAATC	TACTGGATTT	GTTCCCAAAG	AGTATAGCTT	900
50	TGCTAATTAC	ATCACCAATT	CCTTTGGGCA	GTTTGATAAC	TATACGCCCCA	950
	TGCAGTTGGC	TCAGTATGTA	GCAACTATTG	CAAATAATGG	TGTTCTGTGTG	1000
	GCTCCTCGTA	TTGTTGAAGG	CATTTATGGT	AATAATGATA	AGGGAGGACT	1050
	GGGTGACTTG	ATTGAGCAAC	TGCAACCGAC	AGAGATGAAT	AAGGTCAATA	1100
	TATCCGACTC	CGATATGAGC	ATCTTGACAC	AAGGTTTTTA	TCAGGTTGCC	1150
55	CATGGTACTA	GTGGATTGAC	AACCTGGACGT	GCCTTTTCAA	ATGGTGCCCTT	1200
	GGTATCCATT	AGCGGAAAAA	CAGGTACAGC	CGAAAGCTAT	GTGGCAGATG	1250
	GTCAGCAAGC	AACCAATACC	AATGCGGTGG	CCTATGCCCC	ATCTGATAAT	1300
	CCCCAAATCG	CTGTGCGAGT	GGTCTTTTCT	CATAATACCA	ATCTAACAAA	1350
	TGGTGTAGGA	CCTTCCATTG	CGCGTGACAT	TATCAATCTG	TATCAAAAAT	1400
60	ACCATCCAAT	GAAGTAGAAA	GGAAATTA			1428

## 2) INFORMATION FOR SEQ ID NO: 1023

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1442 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-05

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1023

20	TGCCTCTATT	TCAAAGGAGA	TGCCTGGCAT	TAGTATTTCT	ACTTCTTGGG	50
	ATCGAAAGGT	TTTGGAAACT	TCCCTTTCTT	CTATAGTAGG	GAGTGTATCC	100
	AGTGAAAAAG	CTGGTCTCCC	AGCGGAAGAA	GCAGAATCCT	ATCTTAAAAA	150
	AGGCTATTCT	CTAAATGACC	GTGTTGGAAC	CTCCTATTTG	GAAAAGCAAT	200
	ATGAAGAGAC	CTTACAAGGA	AAACGCTCGG	TAAAAGAAAT	CCATCTGGAT	250
25	AAATATGGCA	ATATGGAAAG	CGTGGACACA	ATTGAGGAAG	GTAAGGTAAGG	300
	AAACAATATC	AAACTGACCA	TTGATTTGGC	CTTCCAAGAT	AGCGTGGATG	350
	CTTTGCTGAA	AAGTTATTTT	AATTCCGAGC	TAGGAAATGG	TGGAGCTAAA	400
	TATTCTGAAG	GTGTCTATGC	AGTCGCCCTT	AACCCAAAAA	CAGGTGCTGT	450
	TTTATCCATG	TCAGGGATCA	AACATGACCT	GAAAACGGGA	GAGTTGACTC	500
30	CTGATTCCTT	GGGAACGGTA	ACCAATGTCT	TTGTCCCAGG	TTCGGTTGTC	550
	AAGGCTGCGA	CCATCAGCTC	AGGTGGGAA	AATGGTGTTT	TATCAGGAAA	600
	CCAAACCTTA	ACAGATCAGC	CTATTGTTTT	CCAAGGTTCA	GCTCCAATTT	650
	ATTCTTGGTA	TAAATTGGCA	TATGGATCTT	TTCCTATTAC	AGCTGTGGAA	700
	GCCTTGGAGT	ATTCATCCAA	TGCTTACATG	GTTCAAACCG	CTCTTGGAA	750
35	CATGGGCCAG	ACCTATCAAC	CAAATATGTT	TGTTGGAACC	AGCAATTGGA	800
	AAACAGCTAT	GGGAAAACCT	CGTGCGACCT	TTGGCGAATA	TGGCTTGGGG	850
	GCTGCGACCG	GAATTGACCT	ACCAGATGAA	TCTACTGGAT	TTGTTCCCAA	900
	AGAGTATAGC	TTTGCTAATT	ACATCACCAA	TTCCTTTGGG	CAGTTTGATA	950
	ACTATACGCC	CATGCAGTTG	GCTCAGTATG	TAGCAACTAT	TGCAATAAAT	1000
40	GGTGTTCGTG	TGGCTCCTCG	TATTGTTGAA	GGCATTTATG	GTAATAATGA	1050
	TAAGGGAGGA	CTGGGTGACT	TGATTCAGCA	ACTGCAACCG	ACAGAGATGA	1100
	ATAAGGTCAA	TATATCCGAC	TCCGATATGA	GCATCTTGCA	CCAAGGTTTT	1150
	TATCAGGTTG	CCCATGGTAC	TAGTGGATTG	ACAACCTGGAC	GTGCCTTTTC	1200
	AAATGGTGCC	TTGGTATCCA	TTAGCGGAAA	AACAGGTACA	GCCGAAAGCT	1250
45	ATGTGGCAGA	TGGTCAGCAA	GCAACCAATA	CCAATGCGGT	GGCCTATGCC	1300
	CCATCTGATA	ATCCCCAAAT	CGCTGTGCGA	GTGGTCTTTC	CTCATAATAC	1350
	CAATCTAACA	AATGGTGTAG	GACCTTCCAT	TGCGCGTGAC	ATTATCAATC	1400
	TGTATCAAAA	ATACCATCCA	ATGAAGTAGA	AAGGAAATTA	TG	1442

## 2) INFORMATION FOR SEQ ID NO: 1024

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1445 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-06

5

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1024

	TTGCCTCTAT	TTCAAAGGAG	ATGCCTGGCA	TTAGTATTTT	TACTTCTTGG	50
	GATCGAAAGG	TTTTGGAAAC	TTCCCTTTCT	TCTATAGTAG	GGAGTGTATC	100
10	CAGTGAAAAA	GCTGGTCTCC	CAGCGGAAGA	AGCAGAATCC	TATCTTAAAA	150
	AAGGCTATTC	TCTAAATGAC	CGTGTGGAA	CCTCCTATTT	GGAAAAGCAA	200
	TATGAAGAGA	CCTTACAAGG	AAAACGCTCG	GTAAAAGAAA	TCCATCTGGA	250
	TAAATATGGC	AATATGGAAA	GCGTGGACAC	AATTGAGGAA	GGTAGTAAGG	300
	GAAACAATAT	CAAACTGACC	ATTGATTTGG	CCTTCCAAGA	TAGCGTGGAT	350
15	GCTTTGCTGA	AAAGTTATTT	CAATTCGAG	CTAGGAAATG	GTGGAGCTAA	400
	ATATTCTGAA	GGTGTCTATG	CAGTCGCCCT	TAACCCAAAA	ACAGGTGCTG	450
	TTTTATCCAT	GTCAGGGATC	AAACATGACC	TGAAAACGGG	AGAGTTGACT	500
	CCTGATTCCT	TGGGAACGGT	AACCAATGTC	TTTGTCCCAG	GTTCGGTTGT	550
	CAAGGCTGCG	ACCATCAGCT	CAGGTTGGGA	AAATGGTGTT	TTATCAGGAA	600
20	ACCAAACCTT	AACAGATCAG	CCTATTGTTT	TCCAAGGTTT	AGCTCCAATT	650
	TATTCTTGGT	ATAAATTGGC	ATATGGATCT	TTTCCTATTA	CAGCTGTGGA	700
	AGCCTTGGAG	TATTCATCCA	ATGCTTACAT	GGTTCAAACC	GCTCTTGGA	750
	TCATGGGCCA	GACCTATCAA	CCAAATATGT	TTGTTGGAAC	CAGCAATTTG	800
	GAAACAGCTA	TGGGAAAACT	TCGTGCGACC	TTTGGCGAAT	ATGGCTTGGG	850
25	GGCTGCGACC	GGAATTGACC	TACCAGATGA	ATCTACTGGA	TTTGTTCCTA	900
	AAGAGTATAG	CTTTGCTAAT	TACATCACCA	ATTCCTTTGG	GCAGTTTGAT	950
	AACATATACAC	CCATGCAGTT	GGCTCAGTAT	GTAGCAACTA	TTGCAAATAA	1000
	TGGTGTTTCGT	GTGGCTCCTC	GTATTGTTGA	AGGCATTTAT	GGTAATAATG	1050
	ATAAGGGAGG	ACTGGGTGAC	TTGATTTCAG	AACTGCAACC	GACAGAGATG	1100
30	AATAAGGTCA	ATATATCCGA	TTCCGATATG	AGCATCTTGC	ACCAAGGTTT	1150
	TTATCAGGTT	GCCCATGGTA	CTAGTGGATT	GACAACTGGA	CGTGCCTTTT	1200
	CAAATGGTGC	CTTGGTATCC	ATTAGCGGAA	AAACAGGTAC	AGCCGAAAGC	1250
	TATGTGGCAG	ATGGTCAGCA	AGCAACCAAT	ACCAATGCGG	TGGCCTATGC	1300
	CCCATCTGAT	AATCCCCAAA	TCGCTGTGCG	AGTGGTCTTT	CCTCATAATA	1350
35	CCAATCTAAC	AAATGGTGTA	GGACCTTCCA	TTGCGCGTGA	CATTATCAAT	1400
	CTGTATCAAA	AATACCATCC	AATGAAGTAG	AAAGGAAATT	ATGCT	1445

## 40 2) INFORMATION FOR SEQ ID NO: 1025

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1441 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 50 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-07

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1025

55	TTGCCTCTAT	TTCAAAGGAG	ATGCCTGGCA	TTAGTATTTT	TACTTCTTGG	50
	GATCGAAAGA	TTTTGGAAAC	TTCCCTTTCT	TCTATAGTAG	GGAGTGTATC	100
	CAGTGAAAAA	GCTGGTCTCC	CAGCGGAAGA	AGCAGAATCC	TATCTTAAAA	150
	AAGGCTATTC	TCTAAATGAC	CGTGTGGAA	CCTCGTATTT	GGAAAAGCAA	200
60	TATGAAGAGA	CCTTACAAGG	AAAACGCTCG	GTAAAAGAAA	TCCATCTGGA	250

	TAAATATGGC	AATATGGAAA	GCGTGGACAC	AATTGAGGAA	GGTAGTAAGG	300
	GAAACAATAT	CAAACCTGACC	ATTGATTGG	CCTTCCAAGA	TAGCGTGGAT	350
	GCTTTGCTGA	AAAGTTATTT	CAATTCCGAG	CTAGGAAATG	GTGGAGCCAA	400
	GTATTCTGAG	GGTGTGTATG	CAGTCGCCCT	TAACCCCAAA	ACAGGTGCTG	450
5	TTTTGTCTAT	GTCAGGACTC	AAACATGACC	TGAAAACGGG	AGAGTTGACT	500
	CCTGATTCCCT	TGGGAACGGT	AACCAATGTC	TTTGTCCCAG	GTTTCGGTTGT	550
	CAAGGCTGCG	ACCATCAGCT	CTGGCTGGGA	AAATGGTGT	TTATCAGGAA	600
	ACCAAACCTT	AACAGATCAG	CCTATTGTTT	TCCAAGGTTT	AGCTCCAATT	650
	TATTCTTGGT	ATAAATTGGC	ATATGGATCT	TTTCCTATTA	CAGCTGTGGA	700
10	AGCCTTGGAG	TATTCATCCA	ATGCTTACAT	GGTTCAAACC	GCTCTTGGA	750
	TCATGGGCCA	GACCTATCAA	CCAAATATGT	TTGTTGGAAC	CAGCAATTG	800
	GAAACAGCTA	TGGGAAAAC	TCGTGCGACC	TTTGGCGAAT	ATGGCTTGGG	850
	GGCTGCGACC	GGAATTGACC	TACCAGATGA	ATCTACTGGA	TTTGTTCCCA	900
	AAGACTATAG	CTTTGCTAAT	TACATCACCA	ATGCCTTTGG	GCAGTTTGAT	950
15	AACATACGC	CCATGCAGTT	GGCTCAGTAT	GTAGCAACTA	TTGCAAATGA	1000
	TGGTGTTCGT	GTGGCTCCTC	GTATTGTTGA	AGGCATTTAT	GGTAATAATG	1050
	ATAAGGGAGG	ACTGGGTGAC	TTGATTTCAGC	AACTGCAACC	GACAGAGATG	1100
	AATAAGGTCA	ATATATCCGA	CTCCGATATG	AGTATCTTGC	ACCAAGGATT	1150
	TTACCAAGTA	TCGCATGGAA	CTAGTCCCCT	TACGACAGGA	CGGGCGTTTT	1200
20	CAGATGGCGC	CACTGTTTCT	ATCAGTGGTA	AGACCGGTAC	AGGTGAAAGC	1250
	TATGTAGCTG	GTGGTCAAGA	AGCTAATAAT	ACCAATGCCG	TGGCCTATGC	1300
	TCCAACAGAA	AATCCTCAA	TTGCAGTTGC	AGTAGTCTTT	CCTCATAATA	1350
	CCAATTTAAC	CAAAAATGTT	GGGCCAGCAA	TTGCTCGCGA	CATTATCAAT	1400
	TTATATAACC	AACACCATCC	AATGAATTAG	AAAGGAAGCC	A	1441
25						

## 2) INFORMATION FOR SEQ ID NO: 1026

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1443 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 35 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Streptococcus pneumoniae*
  - (B) STRAIN: StrR-08
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1026

	TATTGCCTCT	ATTTCAAAGG	AGATGCCTGG	CATTAGTATT	TCTACTTCTT	50
45	GGGATCGAAA	GGTTTTGGAA	ACTTCCCTTT	CTTCTATAGT	AGGGAGTGTA	100
	TCCAGTGAAA	AAGCTGGTCT	CCCAGCGGAA	GAAGCAGAAT	CCTATCTTAA	150
	AAAAGGCTAT	TCTCTAAATG	ACCGTGTG	AACCTCCTAT	TTGGAAAAGC	200
	AATATGAAGA	GACCTTACAA	GGAAAACGCT	CGGTAAAAGA	AATCCATCTG	250
	GATAAATATG	GCAATATGGA	AAGCGTGGAC	ACAATTGAGG	AAGGTAGTAA	300
50	GGGAAACAAT	ATCAAACCTGA	CCATTGATTT	GGCCTTCCAA	GATAGCGTGG	350
	ATGCTTTGCT	GAAAAGTTAT	TTCAATTCCG	AGCTAGGAAA	TGGTGGAGCT	400
	AAATATTCTG	AAGGTGTCTA	TGCAGTCGCC	CTTAACCCAA	AAACAGGTGC	450
	TGTTTTATCC	ATGTCAGGGA	TCAAACATGA	CCTGAAAACG	GGAGAGTTGA	500
	CTCCTGATTC	CTTGGGAACG	GTAACCAATG	TCTTTGTCCC	AGGTTCGGTT	550
55	GTCAAGGCTG	CGACCATCAG	CTCAGGTTGG	GAAAATGGTG	TTTTATCAGG	600
	AAACCAAACC	TTAACAGATC	AGCCTATTGT	TTTCCAAGGT	TCAGCTCCAA	650
	TTTATTCTTG	GTATAAATTG	GCATATGGAT	CTTTTCCTAT	TACAGCTGTG	700
	GAAGCCTTGG	AGTATTCATC	CAATGCTTAC	ATGGTTCAAA	CCGCTCTTGG	750
	AATCATGGGC	CAGACCTATC	AACCAAATAT	GTTTGTGGA	ACCAGCAATT	800
60	TGGAAACAGC	TATGGGAAAA	CTTCGTGCGA	CCTTTGGCGA	ATATGGCTTG	850

```

GGGGCTGCGA CCGGAATTGA CCTACCAGAT GAATCTACTG GATTTGTTCC 900
CAAAGAGTAT AGCTTTGCTA ATTACATCAC CAATTCCTTT GGGCAGTTTG 950
ATAACTATAC GCCCATGCAG TTGGCTCAGT ATGTAGCAAC TATTGCAAAT 1000
AATGGTGTTC GTGTGGCTCC TCGTATTGTT GAAGGCATTT ATGGTAATAA 1050
5 TGATAAGGGA GGACTGGGTG ACTTGATTCA GCAACTGCAA CCGACAGAGA 1100
TGAATAAGGT CAATATATCC GACTCCGATA TGAGCATCTT GCACCAAGGT 1150
TTTTATCAGG TTGCCCATGG TACTAGTGGA TTGACAACTG GACGTGCCTT 1200
TTCAAATGGT GCCTTGGTAT CCATTAGCGG AAAAACAGGT ACAGCCGAAA 1250
GCTATGTGGC AGATGGTCAG CAAGCAACCA ATACCAATGC GGTGGCCTAT 1300
10 GCCCCATCTG ATAATCCCCA AATCGCTGTC GCAGTGGTCT TTCCTCATAA 1350
TACCAATCTA ACAAATGGTG TAGGACCTTC CATTGCGCGT GACATTATCA 1400
ATCTGTATCA AAAATACCAT CCAATGAACT AGAAAGGAAA TTA 1443

```

15

## 2) INFORMATION FOR SEQ ID NO: 1027

## (i) SEQUENCE CHARACTERISTICS:

```

20 (A) LENGTH: 1443 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

```

## (ii) MOLECULE TYPE: Genomic DNA

25

## (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Streptococcus pneumoniae
    (B) STRAIN: StrR-09

```

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1027

```

TATTGCCTCT ATTTCAAAGG AGATGCCTGG CATTAGTATT TCTACTTCTT 50
GGGATAGAAA GGTTTTGGAA ACTTCCCTTT CTTCTATAGT TGGGAGTGTA 100
TCCAGTGAAA AAGCTGGTCT CCCAGCGGAA GAAGCAGAAG CCTATCTTAA 150
35 AAAAGGCTAT TCTCTAAATG ATCGTGTAGG AACCTCCTAT TTGGAAAAGC 200
AATATGAAGA GACCTTACAA GGAAAACGCT CGGTAAAAGA AATCCATCTG 250
GATAAATATG GCAATATGGA AAGCGTGGAT ACAATTGAGG AAGGTAGTAA 300
GGGAAACAAT ATCAAACCTGA CCATTGATTT GGCTTTCCAA GATAGCGTGG 350
ATGCTTTTACT GAAAAGTTAT TTCAATTCCG AGCTAGAAAA TGGTGGAGCC 400
40 AAGTATTCTG AAGGTGTCTA TGCAGTCGCC CTTAACCCAA AAACAGGTGC 450
TGTTTTGTCT ATGTCAGGGA TTAACATGA CTTGAAAACG GGAGATTTAA 500
CACCTGATTC CTTGGGAACA GTAACCAATG TCTTTGTCCC GGGTTCTGTT 550
GTCAAGGCGG CGACCATCAG CTCTGGTTGG GAGAATGGAG TCTTATCAGG 600
AAATCAGACC TTGACAGACC AACCATTGTT CTTCCAAGGT TCAGCTCCGA 650
45 TTAATTCTTG GTACACTCAG GCTTACGATT CATTTCCGAT TACAGCGGTG 700
GAAGCCTTGG AGTATTCTTC TAATGCCTAT ATGGTCCAAA CAGCTCTAGG 750
TCTTATGGGG CAGACCTACC AACCCTAATAT GTTTGTCTGGC ACCAGCAATC 800
TAGAGTCTGC TATGGGGAAA TTGCGTTCAA CCTTTGGTGA ATATGGCTTG 850
GGCTCTGCGA CTGGGATTGA CCTACCAGAT GAATCTACTG GATTTGTTCC 900
50 CAAAGAGTAT AGCTTTGCTA ATTACATTAC TAATGCCTTT GGGCAGTTTG 950
ATAACTATAC GCCGATGCAG TTGGCTCAGT ATGTAGCAAC TATTGCAAAT 1000
GATGGTGTTC GTGTGGCTCC TCGTATTGTG GAAGGCATTT ATGGCAATAA 1050
TGATAAGGGA GGACTGGGTG ACTTGATTCA GCAACTGCAA CCGACAGAGA 1100
TGAATAAGGT CAATATATCC GACTCCGATA TGAGTATCTT GCACCAAGGT 1150
55 TTTTATCAGG TTGCTCATGG GACTAGCGGA TTGACAACAG GTCGTGCCTT 1200
TTCCAATGGT GCAGCTGTAT CCATTAGTGG AAAAACAGGT ACCGCCGAAA 1250
GTTATGTAGC AGGTGGCCAA GAAGCCAACA ATACTAATGC TGTAGCCTAT 1300
GCACCATCAG ATAATCCTCA AATAGCTGTT GCTGTTGTCT TCCCTCATAA 1350
CACCAATCTA ACAAATGGTG TAGGACCTTC CATTGCGCGT GATATTATCA 1400
60 ATCTGTATCA AAAATACCAT CCAATGAACT AGAAAGGAAT TTA 1443

```



## 2) INFORMATION FOR SEQ ID NO: 1028

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1446 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1028

20	TATTGCCTCT	ATTTCAAAGG	AGATGCCTGG	CATTAGTATT	TCTACCTCTT	50
	GGGATAGAAA	GGTTTTGGAA	ACTTCTCTTT	CTTCTATAGT	AGGGAGTGTA	100
	TCCAGTGAAA	AAGCTGGTCT	CCCAGCGGAA	GAAGCAGAAG	CCTATCTTAA	150
	AAAAGGCTAT	TCTCTAAATG	ACCGTGTTGG	GACTTCTTAC	CTGGAAAAAC	200
	AATACGAGGA	AACCTTACAA	GGAAAACGCT	CGGTAAAAGA	AATCCATCTG	250
25	GATAAATATG	GCAATATGGA	AAGCGTGGAT	ACAATTGAGG	AAGGTAGTAA	300
	GGGAAACAAT	ATCAAACCTGA	CCATTGATTT	GGCCTTCCAA	GATAGCGTGG	350
	ATGCTTTGCT	GAAAAGTTAT	TTCAATTCCG	AGCTAGGAAA	TGGTGGAGCT	400
	AAGTATTCTG	AAGGCGTGTA	TGCAGTCGCC	CTTAACCCCA	AAACAGGTGC	450
	TGTTTTGTCT	ATGTCAGGAC	TCAAACATGA	CCTGAAAACG	GGAGACTTGA	500
30	CGCCTGATTC	CTTGGGAACG	GTAACCAATG	TCTTTGTCCC	AGGGTCAGTA	550
	GTTAAGGCCG	CTACCATCAG	CTCAGGTTGG	GAAAATGGTG	TTTTATCAGG	600
	AAACCAAACC	TTAACAGATC	AGCCTATTGT	TTTCCAAGGT	TCAGCTCCAA	650
	TTTATTCTTG	GTATAAATTG	GCATATGGAT	CTTTTCCTAT	TACAGCTGTG	700
	GAAGCCTTGG	AGTATTCATC	TAATGCTTAC	ATGGTTCAAA	CCGCTCTTGG	750
35	AATCATGGGC	CAGACCTATC	AACCAAATAT	GTTTGTGGA	ACCAGCAATT	800
	TGGAAACAGC	TATGGGAAAA	CTTCGTGCGA	CCTTTGGCGA	ATATGGCTTG	850
	GGGGCTGCGA	CCGGAATTGA	CCTACCAGAT	GAATCTACTG	GATTGTGTTCC	900
	CAAAGAGTAT	AGCTTTGCTA	ATTACATTAC	TAATGCCTTT	GGGCAGTTTG	950
	ATAACTATAC	GCCGATGCAG	TTGGCTCAGT	ATGTAGCAAC	TATTGCAAAT	1000
40	GATGGTGTTT	GTGTGGCTCC	TCGTATTGTT	GAAGGCATTT	ATGGTAATAA	1050
	TGATAAGGGA	GGACTGGGTG	ACTTGATTCA	GCAACTGCAA	CCGACAGAGA	1100
	TGAATAAGGT	CAATATATCC	GACTCCGATA	TGAGCATCTT	GCACCAAGGT	1150
	TTTTATCAGG	TTGCCCATGG	TACTAGTGGA	TTGACAACTG	GACGTGCCTT	1200
	TTCAAATGGC	GCCTTGGTAT	CCATTAGCGG	AAAAACAGGT	ACAGCCGAAA	1250
45	GCTATGTGGC	AGATGGTCAG	CAAGCAACCA	ATACCAATGC	GGTGGCCTAT	1300
	GCCCCATCTG	ATAATCCCCA	AATCGCTGTT	GCAGTGGTCT	TTCCTCATAA	1350
	TACCAATCTA	ACAAATGGTG	TAGGACCTTC	CATTGCGCGT	GACATTATCA	1400
	ATCTGTATCA	AAAATACCAT	CCAATGAACT	AGAAAGGAAA	TTATGC	1446

## 2) INFORMATION FOR SEQ ID NO: 1029

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1423 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: StrR-11

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1029

```

5      TTGCCTCTAT TTCAAAGGAG ATGCCTGGCA TTAGTATTTT TACTTCTTGG      50
      GATAGAAAGG TTTTGGAAC TTCCCTTTCT TCTATAGTTG GGAGTGTATC      100
10     CAGTGAAAAA GCTGGTCTCC CAGCGGAAGA AGCAGAAGCC TATCTTAAAA      150
      AAGGCTATTC TCTAAATGAC CGTGTAGGAA CCTCCTATTT GGAAAAGCAA      200
      TATGAAGAGA CCTTACAAGG AAAACGCTCG GTAAAAGAAA TCCATCTGGA      250
      TAAATATGGC AATATGGAAA GCGTGGATAC AATTGAGGAA GGTAGTAAGG      300
      GAAACAATAT CAAACTGACC ATTGATTTGG CCTTCCAAGA TAGCGTGGAT      350
15     GCTTTACTGA AAAGTTATTT CAATTCTGAG CTAGAAAATG GTGGAGCCAA      400
      GTATTCTGAA GGTGTCTATG CAGTCGCCCT TAACCCAAAA ACAGGTGCGG      450
      TTTTGTCTAT GTCAGGGATT AAACATGACT TGAAAACGGG AGAGTTGACG      500
      CCTGATTCCT TGGGAACGGT AACCAATGTC TTTGTTCCAG GTTCGGTTGT      550
      CAAGGCGGCG ACCATCAGCT CAGGTTGGGA AAATGGAGTC TTGTCAGGAA      600
20     ACCAGACCTT GACAGACCAG TCCATTGTCT TCCAAGGTTT AGCTCCCATC      650
      AATTCTTGGT ATACTCAGGC TTACGGTTCA TTCCCTATCA CAGCGGTCCA      700
      AGCTCTGGAG TATTCATCCA ATGCTTATAT GGTCCAAACA GCCTTAGGTC      750
      TTATGGGGCA GACCTATCAA CCCAATATGT TTGTCGGCAC CAGCAATCTA      800
      GAGTCTGCTA TGGGTAAATT GCGTTCAACC TTTGCGCAAT ATGGCTTGGG      850
25     GGCTGCGACT GGGATTGATC TACCAGATGA ATCTACTGGA TTTGTTCCCA      900
      AAGACTATAA CTTTGCCAAT TTCATTACCA ATGCCTTTGG GCAGTTTGAT      950
      AACTATACCC CAATGCAATT GGCTCAGTAT GTAGCAACTA TTGCAAATGA      1000
      TGGTGTTCGT GTGGCTCCTC GTATTGTTGA AGGCATTTAT GGTAATAATG      1050
      ATAAGGGAGG ACTGGGTGAC TTGATTCAGC AACTGCAACC GACAGAGATG      1100
30     AATAAGGTCA ATATATCCGA CTCCGATATG AGTGTCTTGC ACCAAGGTTT      1150
      TTATCAGGTT GCTCATGGGA CTAGTGGGTT GACAACTGGC CGTGCCTTTT      1200
      CAAATGGTGC CTTGGTATCC ATTAGCGGAA AAACGGGTAC AGCCGAAAGC      1250
      TATGTGGCAG ATGGTCAGGA AGCAACCAAT ACCAATGCGG TGGCCTATGC      1300
      CCCATCTGAT AATCCCCAAA TCGCTGTGCG AGTGGTCTTT CCTCATAATA      1350
35     CCAATCTAAC AAATGGTGTA GGACCTTCCA TTGCGCGTGA CATTATCAAT      1400
      CTGTATCAAA AATACCATCC AAT      1423

```

## 40 2) INFORMATION FOR SEQ ID NO: 1030

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1447 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

(B) STRAIN: StrR-12

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1030

```

55     TATTGCCTCT ATTTCAAAGG AGATGCCTGG CATTAGTATT TCTACTTCTT      50
      GGGATCGAAA GGTTTTGGAA ACTTCCCTTT CTTCTATAGT AGGGAGTGTA      100
      TCCAGTGAAA AAGCTGGTCT CCCAGCGGAA GAAGCAGAAT CCTATCTTAA      150
      AAAAGGCTAT TCTCTAAATG ACCGTGTTGG AACCTCCTAT TTGGAAAAGC      200
60     AATATGAAGA GACCTTACAA GGAAAACGCT CGGTAAAAGA AATCCATCTG      250

```

	GATAAATATG	GCAATATGGA	AAGCGTGGAC	ACAATTGAGG	AAGGTAGTAA	300
	GGGAAACAAT	ATCAAACCTGA	CCATTGATTT	GGCCTTCCAA	GATAGCGTGG	350
	ATGCTTTGCT	GAAAAGTTAT	TTCAATTCCG	AGCTAGGAAA	TGGTGGAGCT	400
	AAATATTCTG	AAGGTGTCTA	TGCAGTCGCC	CTTAACCCAA	AAACAGGTGC	450
5	TGTTTTATCC	ATGTCAGGGA	TCAAACATGA	CCTGAAAACG	GGAGAGTTGA	500
	CTCCTGATTC	CTTGGGAACG	GTAACCAATG	TCTTTGTCCC	AGGTTCGGTT	550
	GTCAAGGCTG	CGACCATCAG	CTCAGGTTGG	GAAAATGGTG	TTTATCAGG	600
	AAACCAAACC	TTAACAGATC	AGCCTATTGT	TTTCCAAGGT	TCAGCTCCAA	650
	TTTATTCTTG	GTATAAATTG	GCATATGGAT	CTTTTCCTAT	TACAGCTGTG	700
10	GAAGCCTTGG	AGTATTTCATC	CAATGCTTAC	ATGGTTCAAA	CCGCTCTTGG	750
	AATCATGGGC	CAGACCTATC	AACCAAATAT	GTTTGTTGGA	ACCAGCAATT	800
	TGGAAACAGC	TATGGGAAAA	CTTCGTGCGA	CCTTTGGCGA	ATATGGCTTG	850
	GGGGCTGCGA	CCGGAATTGA	CCTACCAGAT	GAATCTACTG	GATTTGTTCC	900
	CAAAGAGTAT	AGCTTTGCTA	ATTACATCAC	CAATTCCTTT	GGGCAGTTTG	950
15	ATAACTATAC	GCCCATGCAG	TTGGCTCAGT	ATGTAGCAAC	TATTGCAAAT	1000
	AATGGTGTTC	GTGTGGCTCC	TCGTATTGTT	GAAGGCATTT	ATGGTAATAA	1050
	TGATAAGGGA	GGACTGGGTG	ACTTGATTCA	GCAACTGCAA	CCGACAGAGA	1100
	TGAATAAGGT	CAATATATCC	GACTCCGATA	TGAGCATCTT	GCACCAAGGT	1150
	TTTTATCAGG	TTGCCCATGG	TACTAGTGGA	TTGACAACCTG	GACGTGCCTT	1200
20	TTCAAATGGT	GCCTTGGTAT	CCATTAGCGG	AAAAACAGGT	ACAGCCGAAA	1250
	GCTATGTGGC	AGATGGTCAG	CAAGCAACCA	ATACCAATGC	GGTGGCCTAT	1300
	GCCCCATCTG	ATAATCCCCA	AATCGCTGTC	GCAGTGGTCT	TTCTTCATAA	1350
	TACCAATCTA	ACAAATGGTG	TAGGACCTTC	CATTGCGCGT	GACATTATCA	1400
	ATCTGTATCA	AAAATACCAT	CCAATGAACT	AGAAAGGAAA	TTATGCT	1447
25						

## 2) INFORMATION FOR SEQ ID NO: 1031

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1446 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-13

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1031

	ATTGCCTCTA	TTTCAAAGGA	GATGCCTGGC	ATTAGTATTT	CTACCTCTTG	50
45	GGATAGAAAG	GTTTTGGAAA	CTTCTCTTTC	TTCTATAGTA	GGGAGTGTAT	100
	CCAGTGAAAA	AGCTGGTCTC	CCAGCGGAAG	AAGCAGAAGC	CTATCTTAAA	150
	AAAGGCTATT	CTCTAAATGA	CCGTGTTGGG	ACTTCTTACC	TGAAAAAACA	200
	ATACGAGGAA	ACCTTACAAG	GAAAACGCTC	GGTAAAAGAA	ATCCATCTGG	250
	ATAAATATGG	CAATATGGAA	AGCGTGGATA	CAATTGAGGA	AGGTAGTAAG	300
50	GGAAACAATA	TCAAACCTGAC	CATTGATTTG	GCCTTCCAAG	ATAGCGTGGA	350
	TGCTTTGCTG	AAAAGTTATT	TCAATTCCGA	GCTAGGAAAT	GGTGGAGCTA	400
	AGTATTCTGA	AGGCGTGTAT	GCAGTCGCCC	TTAACCCCAA	AACAGGTGCT	450
	GTTTGTCTA	TGTCAGGACT	CAAACATGAC	CTGAAAACGG	GAGACTTGAC	500
	GCCTGATTCC	TTGGGAACGG	TAACCAATGT	CTTTGTCCCA	GGGTCAGTAG	550
55	TTAAGGCCGC	TACCATCAGC	TCAGGTTGGG	AAAATGGTGT	TTTATCAGGA	600
	AACCAAACCT	TAACAGATCA	GCCTATTGTT	TTCCAAGGTT	CAGCTCCAAT	650
	TTATTCTTGG	TATAAATTGG	CATATGGATC	TTTTCCTATT	ACAGCTGTGG	700
	AAGCCTTGGA	GTATTCATCT	AATGCTTACA	TGGTTCAAAC	CGCTCTTGGA	750
	ATCATGGGCC	AGACCTATCA	ACCAAATATG	TTTGTTGGAA	CCAGCAATTT	800
60	GGAAACAGCT	ATGGGAAAAC	TTCGTGCGAC	CTTTGGCGAA	TATGGCTTGG	850

	GGGCTGCGAC	CGGAATTGAC	CTACCAGATG	AATCTACTGG	ATTTGTTCCC	900
	AAAGAGTATA	GCTTTGCTAA	TTACATTACT	AATGCCTTTG	GGCAGTTTGA	950
	TAACATACG	CCGATGCAGT	TGGCTCAGTA	TGTAGCAACT	ATTGCAAATG	1000
	ATGGTGTTTCG	TGTGGCTCCT	CGTATTGTTG	AAGGCATTTA	TGGTAATAAT	1050
5	GATAAGGGAG	GACTGGGTGA	CTTGATTTCAG	CAACTGCAAC	CGACAGAGAT	1100
	GAATAAGGTC	AATATATCCG	ACTCCGATAT	GAGCATCTTG	CACCAAGGTT	1150
	TTTATCAGGT	TGCCCATGGT	ACTAGTGGAT	TGACAACTGG	ACGTGCCTTT	1200
	TCAAATGGCG	CCTTGGTATC	CATTAGCGGA	AAAACAGGTA	CAGCCGAAAG	1250
	CTATGTGGCA	GATGGTCAGC	AAGCAACCAA	TACCAATGCG	GTGGCCTATG	1300
10	CCCCATCTGA	TAATCCCCAA	ATCGCTGTTG	CAGTGGTCTT	TCCTCATAAT	1350
	ACCAATCTAA	CAAATGGTGT	AGGACCTTCC	ATTGCGCGTG	ACATTATCAA	1400
	TCTGTATCAA	AAATACCATC	CAATGAACTA	GAAAGGAAAT	TATGCT	1446

15

## 2) INFORMATION FOR SEQ ID NO: 1032

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 1446 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

25

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-14

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1032

	ATTGCCTCTA	TTTCAAAGGA	GATGCCTGGC	ATTAGTATTT	CTACTTCTTG	50
	GGATAGAAAG	GTTTGGGAAA	CTTCCCTTTC	TTCTATAGTT	GGGAGTGTAT	100
	CCAGTGAAAA	AGCTGGTCTC	CCAGCGGAAG	AAGCAGAAGC	CTATCTTAAA	150
35	AAAGGCTATT	CTCTAAATGA	TCGTGTTGGA	ACCTCCTATT	TGGAAAAGCA	200
	ATATGAAGAG	ACCTTACAAG	GGAAACGCTC	GGTAAAAGAA	ATCCATCTGG	250
	ATAAATATGG	CAATATGGAA	AGTGTGGATA	CAATTGAGGA	AGGTAGTAAG	300
	GGAAACAATA	TCAAGCTGAC	CATTGATTTG	GCCTTCCAAG	ATAGCGTGGA	350
	TGCTTTGCTG	AAAAGTTATT	TCAATTCCGA	GCTAGGAAAT	GGTGGAGCCA	400
40	AGTATTCTGA	GGGTGTGTAT	GCAGTCGCCC	TTAACCCCAA	AACAGGTGCT	450
	GTTTTGTCTA	TGTCAGGACT	CAAACATGAC	CTGAAAACGG	GAGAGTTGAC	500
	TCCTGATTCC	TTGGGAACGG	TAACCAATGT	CTTTGTCCCA	GGTTCGGTTG	550
	TCAAGGCTGC	GACCATCAGC	TCTGGCTGGG	AAAATGGAGT	CTTATCAGGA	600
	AACCAGACCT	TGACAGACCA	GTCCATTGTC	TTTCAAGGTT	CAGCTCCCAT	650
45	CAATTCTTGG	TATACTCAGG	CTTACGGTTC	ATTCCCTATC	ACAGCAGTCC	700
	AAGCTCTGGA	GTATTCATCT	AATGCCTATA	TGGTCCAAAC	AGCTTTAGGT	750
	CTTATGGGGC	AGACCTACCA	ACCTAATATG	TTTGTGCGCA	CCAGCAACCT	800
	AGAGTCTGCT	ATGGGGAAAT	TGCGTTCAAC	CTTTGGTGAA	TATGGTTTGG	850
	GTTCTGCGAC	CGGGATTGAC	CTACCAGATG	AATCTACTGG	ATTTGTTCCC	900
50	AAAGACTATA	GCTTTGCTAA	TTACATCACC	AATGCCTTTG	GGCAGTTTGA	950
	TAACATACG	CCGATGCAGT	TGGCTCAGTA	TGTAGCAACT	ATTGCAAATG	1000
	ATGGTGTTTCG	TGTGGCTCCT	CGTATTGTTG	AAGGCATTTA	TGGAAATAAT	1050
	GATAAGGGAG	GCCTAGGCGA	CTTGATTTCAG	CAACTGCAAC	CGACAGAGAT	1100
	GAATAAGGTC	AATATATCCG	ACTCTGATAT	GAGTATTTTG	CACCAAGGTT	1150
55	TTTATCAGGT	TGCTCATGGG	ACTAGTGGAT	TGACAACTGG	ACGTGCCTTT	1200
	TCAAATGGCG	CAGCGGTATC	CATTAGTGGA	AAAACAGGTA	CTGCCGAAAG	1250
	TTATGTTGAG	GGTGGTCAAG	AAGCTAACAA	TACTAATGCT	GTGGCCTATG	1300
	CACCATCAGA	TAATCCTCAA	ATCGCTGTAG	CTGTTGTCTT	CCCTCATAAC	1350
	ACCAACCTTA	CAAATGGTGT	CGGACCTTCC	ATTGCGCGCG	ATATTATCAA	1400
60	CCTCTATAAC	CAACATCATC	CAATGAATTA	GAAAGGAACA	TATGCT	1446

## 2) INFORMATION FOR SEQ ID NO: 1033

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1446 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1033

20	TATTGCCTCT	ATTTCAAAGG	AGATGCCTGG	CATTAGTATT	TCTACTTCTT	50
	GGGATAGAAA	GGTTTTGGAA	ACTTCCCTTT	CTTCTATAGT	TGGGAGTGTA	100
	TCCAGTGAAA	AAGCTGGTCT	CCCAGCGGAA	GAAGCAGAAG	CCTATCTTAA	150
	AAAAGGCTAT	TCTCTAAATG	ACCGTGTTGG	AACCTCGTAT	TTGGAAAAGC	200
	AATATGAAGA	GACCTTACAA	GGAAAACGCT	CGGTAAAAGA	AATCCATCTG	250
25	GATAAATATG	GCAATATGGA	AAGCGTGGAT	ACAATTGAGG	AAGGTAGTAA	300
	GGGAAACAAT	ATCAAGCTGA	CCATTGATTT	GGCCTTCCAA	GATAGCGTGG	350
	ATGCTTTGCT	GAAAAGTTAT	TTCAATTCCG	AGCTAGGAAA	TGGTGGAGCC	400
	AAGTATTCTG	AGGGTGTGTA	TGCAGTCGCC	CTTAACCCCA	AAACAGGTGC	450
	TGTTTTGTCT	ATGTCAGGAC	TCAAACATGA	CCTGAAAACG	GGAGAGTTGA	500
30	CTCCTGATTC	CTTGGGAACG	GTAACCAATG	TCTTTGTCCC	AGGTTTCGGTT	550
	GTTAAGGCCG	CTACCATCAG	CTCAGGTTGG	GAAAATGGTG	TTTTATCAGG	600
	AAACCAAACC	TTAACAGATC	AGCCTATTGT	TTTCCAAGGT	TCAGCTCCAA	650
	TTTATTCTTG	GTATAAATTG	GCATATGGAT	CTTTTCCTAT	TACAGCTGTG	700
	GAAGCCTTGG	AGTATTCATC	CAATGCTTAC	ATGGTTCAAA	CCGCTCTTGG	750
35	AATCATGGGC	CAGACCTATC	AACCAAATAT	GTTTGTGTTG	ACCAGCAATT	800
	TGGAAACAGC	TATGGGAAAA	CTTCGTGCGA	CCTTTGGCGA	ATATGGCTTG	850
	GGGGCTGCGA	CCGGAATTGA	CCTACCAGAT	GAATCTACTG	GATTTGTTCC	900
	CAAAGAGTAT	AGCTTTGCTA	ATTACATCAC	CAATGCCTTT	GGGCAGTTTG	950
	ATAACTATAC	GCCCATGCAG	TTGGCTCAGT	ATGTAGCAAC	TATTGCAAAT	1000
40	GATGGTGTTC	GTGTGGCTCC	TCGTATTGTT	GAAGGCATTT	ATGGTAATAA	1050
	TGATAAGGGA	GGACTGGGTG	ACTTGATTCA	GCAACTGCAA	CCGACAGAGA	1100
	TGAATAAGGT	CAATATATCC	GACTCCGATA	TGAGTATCTT	GCACCAAGGT	1150
	TTTTATCAGG	TTGCTCATGG	GACTAGTGGA	TTGACAACCTG	GACGTGCTTT	1200
	TTCAAATGGT	GCCTTGGTAT	CCATTAGTGG	GAAAACAGGT	ACTGCCGAAA	1250
45	GTTATGTTGC	AGGTGGTCAA	GAAGCCAACA	ATACCAACGC	GGTGGCCTAT	1300
	GCCCCATCAG	ATAATCCTCA	AATCGCTGTT	GCCGTTGTCT	TCCCTCATAA	1350
	CACCAATCTA	ACAAATGGTG	TTGGACCTTC	TATTGCACGC	GATATTATCA	1400
	ACCTCTATAA	CCAACACCAT	CCAATGAATT	AGAAAGGAAC	TTATGC	1446

## 2) INFORMATION FOR SEQ ID NO: 1034

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1670 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

782

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-01

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1034

```

5      GGAAACTGCA GAGGTCAAGG GGATTGATTT TACAACCAGT CCCAATCGTA      50
      GTTATCCAAA CGGACAATTT GCTTCTAGTT TTATCGGACT AGCTCAGCTC      100
10     CATGAAAATG AAGATGGCAG CAAGAGCTTA CTGGGAACCT CTGGAATGGA      150
      GAGTTCCTTG AACAGTATTC TTGCAGGGAC AGACGGTATT ATTACCTATG      200
      AAAAAGACCG TGTAGGAAAT ATCGTACCAG GTACAGAACT GGTATCGCAA      250
      CAAACTGTGG ATGGCAAGGA TGTTTATACA ACATTGTCTA GTCCGCTACA      300
      ATCTTTCATG GAAACTCAGA TGGATGCCTT TCTAGAAAAA GTAAAAGGTA      350
15     AGTATATGAC CGCGACCTTG GTCAGTGCAA AGACCGGTGA AATCCTCGCT      400
      ACCACCCAAC GACCTACCTT TAATGCAGAT ACTAAAGAAG GAATCACTGA      450
      GGACTTTGTT TGGCGTGATA TTCTTTATCA AAGTAACTAT GAACCAGGAT      500
      CAGCCATGAA GGTATGACG TTAGCTTCTT CTATTGATAA TAATACCTTC      550
      CCAAGTGGAG AATACTTCAA TAGCAGTGAA TTCAAATAG CGGATGCGAC      600
20     GACTCGAGAT TGGGATGTTA ATGATGGTTT GACTACTGGT GGGATGATGA      650
      CTTTCTTACA AGGTTTCGCT CACTCCAGTA ATGTTGGAAT GAGTCTACTT      700
      GAACAAAAAA TGGGAGATGC TACTTGGTTG GATTATCTAA AACGCTTTAA      750
      ATTTGGGGTT CCAACTCGCT TTGGCTTGAC AGATGAATAC GCTGGTCAAC      800
      TTCCAGCTGA TAATATTGTT AGTATTGCTC AAAGCTCATT TGGGCAAGGA      850
25     ATTTCAGTGA CACAAACACA AATGCTTCGT GCCTTTACAG CTATTGCTAA      900
      TGATGGAGTT ATGCTGGAGC CAAAATTTAT AAGTGCTATT TATGATACTA      950
      ACAATCAGTC TGTACGTAAG TCACAAAAAG AAATAGTAGG AAATCCTGTT      1000
      TCCAAAGAGG CAGCAAGCAC AACTCGAAAT CACATGATCT TAGTTGGGAC      1050
      GGACCCTCTA TATGGAAC TAATAATCA CTACACAGGA AAGCCAATTA      1100
30     TAACAGTTCC TGGACAAAAT GTAGCAGTTA AATCCGGTAC GGCTCAAATC      1150
      GCTGATGAGA AAAATGGAGG ATACTTGCTT GGTCTACCA ATTATATTTT      1200
      CTCAGTTGTG ACTATGAATC CTGCTGAAAA TCCTGATTTT ATCTTGATG      1250
      TAACGGTTCA ACAGCCTGAG CATTATTCAG GTATCCAGTT GGGAGAATTT      1300
      GCCACCCCAA TCTTGGAGCG GGCTTCAGCT ATGAAAGAAT CTCTCAATCT      1350
35     TCAATCTCCA GCCAAAATT TAGATAAAGT TACGACAGAA TCTTCTTATG      1400
      CAATGCCTAG CATCAAGGAT ATTTACCTG GTGAGTTGGC GGAAGCCTTA      1450
      CGCCGAAATA TTGTGCAACC AATCGTTGTA GGTACTGGAA CAAAGATTAA      1500
      AGAGACTTCT GTAGAAGAAG GGACCAATCT TGCACCAAAC CAACAAGTTC      1550
      TCCTTTTATC GGATAAGGTA GAAGAAATTC CAGACATGTA TGGCTGGAAA      1600
40     AAAGAGACTG CCGAGACCTT TGCTAAATGG TTGGATATTG AACTGGAATT      1650
      TGAAGGTTCA GGTTCCGTTG      1670

```

## 45 2) INFORMATION FOR SEQ ID NO: 1035

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1683 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-02

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1035

	AAAGAATTGG	AAACTGCAGA	GGTCAAGGGG	ATTGATTTTA	CAACCAGTCC	50
	CAATCGTAGT	TACCCAAATG	GACAATTTGC	TTCTAGTTTT	ATCGGTCTAG	100
	CTCAGCTCCA	TGAAAATGAA	GATGGAAGCA	AGAGTTTGCT	GGGAACCTCT	150
	GGAATGGAGA	GTTCCTTGAA	CAGTATTCTT	GCAGGGACAG	ACGGCATTAT	200
5	TACCTATGAA	AAGGATCGTC	TGGGCAATAT	TGTACCCGGA	ACAGAACAAG	250
	TTTCCCAACA	AACGGTGGAT	GGCAAGGATG	TTTACACAAC	CATTTCCAGC	300
	CCCCTCCAGT	CCTTCATGGA	AACTCAGATG	GATGCCTTTC	TAGAAAAAGT	350
	AAAAGGTAAG	TATATGACCG	CGACCTTGGT	CAGTGCAAAG	ACCGGTGAAA	400
	TCCTCGCTAC	CACCCAACGA	CCTACCTTTA	ATGCAGATAC	TAAAGAAGGA	450
10	ATCACTGAGG	ACTTTGTTTG	GCGTGATATT	CTTTATCAAA	GTAACATGA	500
	ACCAGGATCA	GCCATGAAGG	TTATGACGTT	AGCTTCTTCT	ATTGATAATA	550
	ATACCTTCCC	AAGTGGAGAA	TACTTCAATA	GCAGTGAATT	CAAAATAGCG	600
	GATGCGACGA	CTCGAGATTG	GGATGTTAAT	GATGGTTTGA	CTACTGGTGG	650
	GATGATGACT	TTCTTACAAG	GTTTCGCTCA	CTCCAGTAAT	GTTGGAATGA	700
15	GTCTACTTGA	ACAAAAAATG	GGAGATGCTA	CTTG GTTGA	TTATCTAAAA	750
	CGCTTTAAAT	TTGGGGTTCC	AACTCGCTTT	GGCTTGACAG	ATGAATACGC	800
	TGGTCAACTT	CCAGCTGATA	ATATTGTTAG	TATTGCTCAA	AGCTCATTTG	850
	GGCAAGGAAT	TTCAGTGACA	CAAACACAAA	TGCTTCGTGC	CTTTACAGCT	900
	ATTGCTAATG	ATGGAGTTAT	GCTGGAGCCA	AAATTTATAA	GTGCTATTTA	950
20	TGATACTAAC	AATCAGTCTG	TACGTAAGTC	ACAAAAAGAA	ATAGTAGGAA	1000
	ATCCTGTTTC	CAAAGAGGCA	GCAAGCACAA	CTCGAAATCA	CATGATCTTA	1050
	GTTGGGACGG	ACCCTCTATA	TGGAACATATG	TATAATCACT	ACACAGGAAA	1100
	GCCAATTATA	ACAGTTCCTG	GACAAAATGT	AGCAGTTAAA	TCCGGTACGG	1150
	CTCAAATCGC	TGATGAGAAA	AATGGAGGAT	ACTTGGTTGG	TTCTACCAAT	1200
25	TATATTTTCT	CAGTTGTGAC	TATGAATCCT	GCTGAAAATC	CTGATTTTAT	1250
	CTTGATATGA	ACGGTTCAAC	AGCCTGAGCA	TTATTTCAGGT	ATCCAGTTGG	1300
	GAGAATTTGC	CACCCCAATC	TTGGAGCGGG	CTTCAGCTAT	GAAAGAATCT	1350
	CTCAATCTTC	AATCTCCAGC	CAAAAATTTA	GATAAAGTTA	CGACAGAATC	1400
	TTCTTATGCA	ATGCCTAGCA	TCAAGGATAT	TTCACCTGGT	GAGTTGGCGG	1450
30	AAGCCTTACG	CCGAAATATT	GTGCAACCAA	TCGTTGTAGG	TACTGGAACA	1500
	AAGATTAAAG	AGACTTCTGT	AGAAGAAGGG	ACCAATCTTG	CACCAAACCA	1550
	ACAAGTTCTC	CTTTTATCGG	ATAAGGTAGA	AGAAATTCCA	GACATGTATG	1600
	GCTGGAAAAA	AGAGACTGCC	GAGACCTTTG	CTAAATGGTT	GGATATTGAA	1650
	CTGGAATTTG	AAGGTTTCAGG	TTCCGTTGTT	CAG		1683
35						

## 2) INFORMATION FOR SEQ ID NO: 1036

- 40 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1682 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 45 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Streptococcus pneumoniae*
  - (B) STRAIN: StrR-03
- 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1036

	TCAAAAAAGA	ATTGGAAACT	GCAGAGGTCA	AGGGGATTGA	TTTTACAACC	50
55	AGTCCCAATC	GTAGTTACCC	AAATGGACAA	TTTGCTTCTA	GTTTTATCGG	100
	TCTAGCTCAG	CTCCATGAAA	ATGAAGATGG	AAGCAAGAGT	TTGCTGGGAA	150
	CCTCTGGAAT	GGAGAGTTCC	TTGAACAGTA	TTCTTGCAGG	GACAGACGGC	200
	ATTATTACCT	ATGAAAAGGA	TCGTCTGGGC	AATATTGTAC	CCGGAACAGA	250
	ACAAGTTTCC	CAACAAACGG	TGGATGGCAA	GGATGTTTAC	ACAACCATTT	300
60	CCAGCCCCCT	CCAGTCCTTC	ATGGAAACTC	AGATGGATGC	CTTCTAGAA	350

	AAAGTAAAAG	GTAAGTATAT	GACCGCGACC	TTGGTCAGTG	CAAAGACCGG	400
	TGAAATCCTC	GCTACCACCC	AACGACCTAC	CTTTAATGCA	GATACTAAAG	450
	AAGGAATCAC	TGAGGACTTT	GTTTGGCGTG	ATATTCTTTA	TCAAAGTAAC	500
	TATGAACCAG	GATCAGCCAT	GAAGGTTATG	ACGTTAGCTT	CTTCTATTGA	550
5	TAATAATACC	TTCCCAAGTG	GAGAATACTT	CAATAGCAGT	GAATTCAAAA	600
	TAGCGGATGC	GACGACTCGA	GATTGGGATG	TTAATGATGG	TTTGACTACT	650
	GGTGGGATGA	TGACTTTCTT	ACAAGGTTTC	GCTCACTCCA	GTAATGTTGG	700
	AATGAGTCTA	CTTGAACAAA	AAATGGGAGA	TGCTACTTGG	TTGGATTATC	750
	TAAAACGCTT	TAAATTTGGG	GTTCCAACCTC	GCTTTGGCTT	GACAGATGAA	800
10	TACGCTGGTC	AACTTCCAGC	TGATAATATT	GTTAGTATTG	CTCAAAGCTC	850
	ATTTGGGCAA	GGAATTTTCAG	TGACACAAAC	ACAAATGCTT	CGTGCCCTTA	900
	CAGCTATTGC	TAATGATGGA	GTTATGCTGG	AGCCAAAATT	TATAAGTGCT	950
	ATTTATGATA	CTAACAATCA	GTCTGTACGT	AAGTCACAAA	AAGAAATAGT	1000
	AGGAAATCCT	GTTTCCAAAG	AGGCAGCAAG	CACAACTCGA	AATCACATGA	1050
15	TCTTAGTTGG	GACGGACCCT	CTATATGGAA	CTATGTATAA	TCACTACACA	1100
	GGAAAGCCAA	TTATAACAGT	TCCTGGACAA	AATGTAGCAG	TTAAATCCGG	1150
	TACGGCTCAA	ATCGCTGATG	AGAAAAATGG	AGGATACTTG	GTTGGTTCTA	1200
	CCAATTATAT	TTTCTCAGTT	GTGACTATGA	ATCCTGCTGA	AAATCCTGAT	1250
	TTTATCTTGT	ATGTAACGGT	TCAACAGCCT	GAGCATTATT	CAGGTATCCA	1300
20	GTTGGGAGAA	TTTGCCACCC	CAATCTTGGA	GCGGGCTTCA	GCTATGAAAG	1350
	AATCTCTCAA	TCTTCAATCT	CCAGCCAAAA	ATTTAGATAA	AGTTACGACA	1400
	GAATCTTCTT	ATGCAATGCC	TAGCATCAAG	GATATTTTAC	CTGGTGAGTT	1450
	GGCGGAAGCC	TTGACCCGAA	ATATTGTGCA	ACCAATCGTT	GTAGGTACTG	1500
	GAACAAAGAT	TAAAGAGACT	TCTGTAGAAG	AAGGGACCAA	TCTTGCACCA	1550
25	AACCAACAAG	TTCTCCTTTT	ATCGGATAAG	GTAGAAGAAA	TTCCAGACAT	1600
	GTATGGCTGG	AAAAAAGAGA	CTGCCGAGAC	CTTTGCTAAA	TGGTTGGATA	1650
	TTGAACTGGA	ATTTGAAGGT	TCAGGTTCCG	TT		1682

30

## 2) INFORMATION FOR SEQ ID NO: 1037

## (i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 1681 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

40

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-04

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1037

	CAAAAAAGAA	TTGGAAACTG	CAGAGGTCAA	GGGGATTGAT	TTTACAACCA	50
	GTCCCAATCG	TAGTTATCCA	AACGGACAAT	TTGCTTCTAG	TTTTATCGGA	100
	CTAGCTCAGC	TCCATGAAA	TGAAGATGGC	AGCAAGAGCT	TACTGGGAAC	150
50	CTCTGGAATG	GAGAGTTCCT	TGAACAGTAT	TCTTGCAGGG	ACAGACGGTA	200
	TTATTACCTA	TGAAAAAGAC	CGTGTAGGAA	ATATCGTACC	AGGTACAGAA	250
	CTGGTATCGC	AACAAACTGT	GGATGGCAAG	GATGTTTATA	CAACATTGTC	300
	TAGTCCGCTA	CAATCTTTCA	TGGAAACTCA	GATGGATGCC	TTTCTAGAAA	350
	AAGTAAAAGG	TAAGTATATG	ACCGCGACCT	TGGTCAGTGC	AAAGACCGGT	400
55	GAAATCCTCG	CTACCACCCA	ACGACCTACC	TTTAATGCAG	ATACTAAAGA	450
	AGGAATCACT	GAGGACTTTG	TTTGGCGTGA	TATTCTTTAT	CAAAGTAACT	500
	ATGAACCAGG	ATCAGCCATG	AAGGTTATGA	CGTTAGCTTC	TTCTATTGAT	550
	AATAATACCT	TCCCAAGTGG	AGAATACTTC	AATAGCAGTG	AATTCAAAAT	600
	AGCGGATGCG	ACGACTCGAG	ATTGGGATGT	TAATGATGGT	TTGACTACTG	650
60	GTGGGATGAT	GACTTTCTTA	CAAGGTTTCG	CTCACTCCAG	TAATGTTGGA	700



	ATGAGTCTAC	TTGAACAAAA	AATGGGAGAT	GCTACTTGGT	TGGATTATCT	750
	AAAACGCTTT	AAATTTGGGG	TTCCAACCTG	CTTTGGCTTG	ACAGATGAAT	800
	ACGCTGGTCA	ACTTCCAGCT	GATAATATTG	TTAGTATTGC	TCAAAGCTCA	850
	TTTGGGCAAG	GAATTTTCAGT	GACACAAACA	CAAATGCTTC	GTGCCTTTAC	900
5	AGCTATTGCT	AATGATGGAG	TTATGCTGGA	GCCAAAATTT	ATAAGTGCTA	950
	TTTATGATAC	TAACAATCAG	TCTGTACGTA	AGTCACAAAA	AGAAATAGTA	1000
	GGAAATCCTG	TTTCCAAAGA	GGCAGCAAGC	ACAACCTCGA	ATCACATGAT	1050
	CTTAGTTGGG	ACGGACCCCTC	TATATGGAAC	TATGTATAAT	CACTACACAG	1100
	GAAAGCCAAT	TATAACAGTT	CCTGGACAAA	ATGTAGCAGT	TAAATCCGGT	1150
10	ACGGCTCAAA	TCGCTGATGA	GAAAAATGGA	GGATACTTGG	TTGGTTCTAC	1200
	CAATTATATT	TTCTCAGTTG	TGACTATGAA	TCCTGCTGAA	AATCCTGATT	1250
	TTATCTTGTA	TGTAACGGTT	CAACAGCCTG	AGCATTATTC	AGGTATCCAG	1300
	TTGGGAGAAT	TTGCCACCCC	AATCTTGGAG	CGGGCTTCAG	CTATGAAAGA	1350
	ATCTCTCAAT	CTTCAATCTC	CAGCCAAAAA	TTAGATAAAA	GTTACGACAG	1400
15	AATCTTCTTA	TGCAATGCCT	AGCATCAAGG	ATATTTCAAC	TGGTGAGTTG	1450
	GCGGAAGCCT	TACGCCGAAA	TATTGTGCAA	CCAATCGTTG	TAGGTACTGG	1500
	AACAAAGATT	AAAGAGACTT	CTGTAGAAGA	AGGGACCAAT	CTTGCACCAA	1550
	ACCAACAAGT	TCTCCTTTTA	TCGGATAAGG	TAGAAGAAAT	TCCAGACATG	1600
	TATGGCTGGA	AAAAAGAGAC	TGCCGAGACC	TTTGCTAAAT	GGTTGGATAT	1650
20	TGAACCTGGA	TTTGAAGGTT	CAGGTTCCGT	T		1681

## 2) INFORMATION FOR SEQ ID NO: 1038

25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1685 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

30

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-05

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1038

40	AAAAAAGAAT	TGGAAACTGC	AGAGGTCAAG	GGGATTGATT	TTACAACCAG	50
	TCCAATCGT	AGTTATCCAA	ACGGACAATT	TGCTTCTAGT	TTTATCGGAC	100
	TAGCTCAGCT	CCATGAAAAT	GAAGATGGCA	GCAAGAGCTT	ACTGGGAACC	150
	TCTGGAATGG	AGAGTTCCTT	GAACAGTATT	CTTGCAGGGA	CAGACGGTAT	200
	TATTACCTAT	GAAAAAGACC	GTGTAGGAAA	TATCGTACCA	GGTACAGAAC	250
45	TGGTATCGCA	ACAAACTGTG	GATGGCAAGG	ATGTTTATAC	AACATTGTCT	300
	AGTCCGCTAC	AATCTTTCAT	GGAAACTCAG	ATGGATGCCT	TTCTAGAAAA	350
	AGTAAAAGGT	AAGTATATGA	CCGCGACCTT	GGTCAGTGCA	AAGACCGGTG	400
	AAATCCTCGC	TACCACCCAA	CGACCTACCT	TTAATGCAGA	TACTAAAGAA	450
	GGAATCACTG	AGGACTTTGT	TTGGCGTGAT	ATTCTTTATC	AAAGTAACTA	500
50	TGAACCAGGA	TCAGCCATGA	AGGTTATGAC	GTTAGCTTCT	TCTATTGATA	550
	ATAATACCTT	CCCAAGTGGA	GAATACTTCA	ATAGCAGTGA	ATTCAAAATA	600
	GCGGATGCGA	CGACTCGAGA	TTGGGATGTT	AATGATGGTT	TGACTACTGG	650
	TGGGATGATG	ACTTTCTTAC	AAGGTTTCGC	TCACTCCAGT	AATGTTGGAA	700
	TGAGTCTACT	TGAACAAAAA	ATGGGAGATG	CTACTTGGTT	GGATTATCTA	750
55	AAACGCTTTA	AATTTGGGGT	TCCAACCTCGC	TTTGGCTTGA	CAGATGAATA	800
	CGCTGGTCAA	CTTCCAGCTG	ATAATATTGT	TAGTATTGCT	CAAAGCTCAT	850
	TTGGGCAAGG	AATTTTCAGT	ACACAAACAC	AAATGCTTCG	TGCCTTTACA	900
	GCTATTGCTA	TATGATGGAGT	TATGCTGGAG	CCAAAATTTA	TAAGTGCTAT	950
	TTATGATACT	AACAATCAGT	CTGTACGTAA	GTCACAAAAA	GAAATAGTAG	1000
60	GAAATCCTGT	TTCCAAAGAG	GCAGCAAGCA	CAACTCGAAA	TCACATGATC	1050

```

TTAGTTGGGA CGGACCCTCT ATATGGAAC TGTGATAATC ACTACACAGG 1100
AAAGCCAATT ATAACAGTTC CTGGACAAAA TGTAGCAGTT AAATCCGGTA 1150
CGGCTCAAAAT CGCTGATGAG AAAAATGGAG GATACTTGGT TGGTTCTACC 1200
AATTATATTT TCTCAGTTGT GACTATGAAT CCTGCTGAAA ATCCTGATTT 1250
5 TATCTTGTAT GTAACGGTTC AACAGCCTGA GCATTATTCA GGTATCCAGT 1300
TGGGAGAATT TGCCACCCCA ATCTTGGAGC GGGCTTCAGC TATGAAAGAA 1350
TCTCTCAATC TTCAATCTCC AGCCAAAAAT TTAGATAAAG TTACGACAGA 1400
ATCTTCTTAT GCAATGCCTA GCATCAAGGA TATTTACCT GGTGAGTTGG 1450
CGGAAGCCTT ACGCCGAAAT ATGTGCAAC CAATCGTTGT AGGTACTGGA 1500
10 ACAAAGATTA AAGAGACTTC TGTAAGAGAA GGGACCAATC TTGCACCAAA 1550
CCAACAAGTT CTCCTTTTAT CGGATAAGGT AGAAGAAAT CCAGACATGT 1600
ATGGCTGGAA AAAAGAGACT GCCGAGACCT TTGCTAAATG GTTGGATATT 1650
GAACTGGAAT TTGAAGGTT AGGTTCCGTT GTTCA 1685

```

15

## 2) INFORMATION FOR SEQ ID NO: 1039

```

(i) SEQUENCE CHARACTERISTICS:
20 (A) LENGTH: 1679 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Streptococcus pneumoniae
    (B) STRAIN: StrR-06

```

30

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1039

```

AAAAGAATTG GAAACTGCAG AGGTCAAGGG GATTGATTTT ACAACCAGTC 50
CCAATCGTAG TTATCCAAAC GGACAATTTG CTTCTAGTTT TATCGGACTA 100
35 GCTCAGCTCC ATGAAAATGA AGATGGCAGC AAGAGCTTAC TGGGAACCTC 150
TGGAATGGAG AGTTCCTTGA ACAGTATTCT TGCAGGGACA GACGGTATTA 200
TTACCTATGA AAAAGACCGT GTAGGAAATA TCGTACCAGG TACAGAACTG 250
GTATCGCAAC AAAGTGTGGA TGGCAAGGAT GTTTATACAA CATTGTCTAG 300
TCCGCTACAA TCTTTCATGG AAAGTCAGAT GGATGCCTTT CTAGAAAAAG 350
40 TAAAAGGTAA GTATATGACC GCGACCTTGG TCAGTGCAA GACCGGTGAA 400
AATCTCGCTA CCACCAACG ACCTACCTTT AATGCAGATA CTAAAGAAGG 450
AATCACTGAG GACTTTGTTT GCGGTGATAT TCTTTATCAA AGTAACTATG 500
AACCAGGATC AGCCATGAAG GTTATGACGT TAGCTTCTTC TATTGATAAT 550
AATACCTTCC CAAGTGGAGA ATACTTCAAT AGCAGTGAAT TCAAAATAGC 600
45 GGATGCGACG ACTCGAGATT GGGATGTTAA TGATGGTTTG ACTACTGGTG 650
GGATGATGAC TTTCTTACAA GGTTTCGCTC ACTCCAGTAA TGTGGAATG 700
AGTCTACTTG AACAAAAAAT GGGAGATGCT ACTTGGTTGG ATTATCTAAA 750
ACGCTTTAA TTTGGGGTTC CAAGTCGCTT TGGCTTGACA GATGAATACG 800
CTGGTCAACT TCCAGCTGAT AATATTGTTA GTATTGCTCA AAGCTCATTT 850
50 GGGCAAGGAA TTTCAGTGAC ACAAACACAA ATGCTTCGTG CCTTTACAGC 900
TATTGCTAAT GATGGAGTTA TGCTGGAGCC AAAATTTATA AGTGCTATTT 950
ATGATACTAA CAATCAGTCT GTACGTAAGT CACAAAAAGA AATAGTAGGA 1000
AATCCTGTTT CCAAAGAGGC AGCAAGCACA ACTCGAAATC ACATGATCTT 1050
AGTTGGGACG GACCCTCTAT ATGGAACAT GTATAATCAC TACACAGGAA 1100
55 AGCCAATTAT AACAGTTCTT GGACAAAATG TAGCAGTTAA ATCCGGTACG 1150
GCTCAAAATCG CTGATGAGAA AAATGGAGGA TACTGGTTTG GTTCTACCAA 1200
TTATATTTTC TCAGTTGTGA CTATGAATCC TGCTGAAAAT CCTGATTTTA 1250
TCTTGATATG AACGGTTCAA CAGCCTGAGC ATTATTCAGG TATCCAGTTG 1300
GGAGAATTTG CCACCCCAAT CTTGGAGCGG GCTTCAGCTA TGAAAGAATC 1350
60 TCTCAATCTT CAATCTCCAG CAAAAAATTT AGATAAAGTT ACGACAGAAT 1400

```

	CTTCTTATGC	AATGCCTAGC	ATCAAGGATA	TTTCACCTGG	TGAGTTGGCG	1450
	GAAGCCTTAC	GCCGAAATAT	TGTGCAACCA	ATCGTTGTAG	GTACTGGAAC	1500
	AAAGATTAAA	GAGACTTCTG	TAGAAAGAAG	GACCAATCTT	GCACCAAACC	1550
	AACAAGTTCT	CCTTTTATCG	GATAAGGTAG	AAGAAATTCC	AGACATGTAT	1600
5	GGCTGGAAAA	AAGAGACTGC	CGAGACCTTT	GCTAAATGGT	TGGATATTGA	1650
	ACTGGAATTT	GAAGGTTTCA	GTTCCGTTG			1679

## 10 2) INFORMATION FOR SEQ ID NO: 1040

## (i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH: 1685 bases
	(B) TYPE: Nucleic acid
15	(C) STRANDEDNESS: Double
	(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 20 (vi) ORIGINAL SOURCE:

	(A) ORGANISM: <i>Streptococcus pneumoniae</i>
	(B) STRAIN: StrR-07

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1040

25	TAAAAAGGAT	CTAAAAGACG	CTAGTGTTGA	AGGAATTGAC	TTCACAACCTA	50
	GCCCTAATAG	AAGCTATCCA	AATGGACAAT	TCGCTTCTAG	TTTTATTGGT	100
	TTGGCCCCAAC	TCCATGAAAA	TGAGGATGGT	AGCAAGAGTT	TGCTGGGAAC	150
	TTCTGGGATG	GAGAGTTCCT	TGAACAGTAT	TCTTGCAGGG	AAAGACGGTA	200
30	TTATTACCTA	TGAAAAAGAT	CGTCTGGGTA	ATATTGTCCC	TGGAACAGAA	250
	CAAGTTTCCC	AACAAACGGT	AGATGGCAAG	GATGTTTATA	CAACATTGTC	300
	TAGTCCGCTA	CAATCTTTCA	TGGAAACTCA	GATGGATGCC	TTTCTAGAAA	350
	AAGTAAAAGG	TAAGTATATG	ACCGCGACCT	TGGTCAGTGC	AAAGACCGGT	400
	GAAATCCTCG	CTACCACCCA	ACGACCTACC	TTTAATGCAG	ATACTAAAGA	450
35	AGGAATCACT	GAGGACTTTG	TTTGGCGTGA	TATTCTTTAT	CAAAGTAACT	500
	ATGAACCAGG	ATCAGCCATG	AAGGTTATGA	CGTTAGCTTC	TTCTATTGAT	550
	AATAATACCT	TCCCAAGTGG	AGAATACTTC	AATAGCAGTG	AATTCAAAAT	600
	AGCGGATCGG	ACGACTCGAG	ATTGGGATGT	TAATGAAGGT	TTGACTACTG	650
	GTGGGATGAT	GACTTTCTTA	CAAGGTTTCG	CTCACTCCAG	TAATGTTGGA	700
40	ATGAGTCTAC	TTGAACAAAA	AATGGGAGAT	GCTACTTGGT	TGGATTATCT	750
	AAAACGCTTT	AAATTTGGGG	TTCCAACCTC	CTTTGGCTTG	ACAGATGAAT	800
	ACGCTGGTCA	ACTTCCAGCT	GATAATATTG	TTAGTATTGC	TCAAAGCTCA	850
	TTTGGGCAAG	GAATTTTCAGT	GACACAAACA	CAAATGCTTC	GTGCCTTTAC	900
	AGCTATTGCT	AATGATGGAG	TTATGCTGGA	GCCAAAATTT	ATAAGTGCTA	950
45	TTTATGATAC	TAACAATCAG	TCTGTACGTA	AGTCACAAAA	AGAAATAGTA	1000
	GGAAATCCTG	TTTCCAAAGA	GGCAGCAAGC	ACAACCTCGAA	ATCACATGAT	1050
	CTTAGTTGGG	ACGGACCCTC	TATATGGAAC	TATGTATAAT	CACTACACAG	1100
	GAAAGCCAAT	TATAACAGTT	CCTGGACAAA	ATGTAGCAGT	TAAATCCGGT	1150
	ACGGCTCAAA	TCGCTGATGA	GAAAAATGGA	GGATACTTGG	TTGGTTCTAC	1200
50	CAATTATATT	TTCTCAGCTG	TGACTATGAA	TCCTGCTGAA	AATCCTGATT	1250
	TTATCTTGTA	TGTAACGGTT	CAACAGCCTG	AGCATTATTC	AGGTATCCAG	1300
	TTGGGAGAAT	TTGCCACCCC	AATCTTGGAG	CGGGCTTCAG	CTATGAAAGA	1350
	ATCTCTCAAT	CTTCAATCTC	CAGCCAAAAA	TTTAGATAAA	GTTACGACAG	1400
	AATCTTCTTA	TGCAATGCCT	AGCATCAAGG	ATATTTTACC	TGGTGAGTTG	1450
55	GCGGAAGCCT	TACGCCGAAA	TATTGTGCAA	CCAATCGTTG	TAGGTACTGG	1500
	AACAAAGATT	AAAGAGACTT	CTGTAGAAGA	AGGGACCAAT	CTTGCAACAA	1550
	ACCAACAAGT	TCTCCTTTTA	TCGGATAAAG	TAGAAGAAAT	TCCAGACATG	1600
	TATGGCTGGA	AAAAAGAGAC	TGCCGAGACC	TTTGCTAAAT	GGTTGGATAT	1650
60	TGAACTGGAA	TTTGAAGGTT	CAGGTTCCGT	CGTTC		1685

## 2) INFORMATION FOR SEQ ID NO: 1041

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1696 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-08

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1041

20	GTCTATCAAA	AAAGAATTGG	AAACTGCAGA	GGTCAAGGGG	ATTGATTTTA	50
	CAACCACTCC	CAATCGTAGT	TATCCAAACG	GACAATTTGC	TTCTAGTTTT	100
	ATCGGACTAG	CTCAGACTCCA	TGAAAATGAA	GATGGCAGCA	AGAGCTTACT	150
	GGGAACCTCT	GGAATGGAGA	GTTCCCTTGA	CAGTATTCTT	GCAGGGACAG	200
	ACGGTATTAT	TACCTATGAA	AAAGACCGTG	TAGGAAATAT	CGTACCAGGT	250
	ACAGAACTGG	TATCGCAACA	AACTGTGGAT	GGCAAGGATG	TTTATACAAC	300
25	ATTGTCCTAGT	CCGCTACAAT	CTTTCATGGA	AACTCAGATG	GATGCCTTTC	350
	TAGAAAAAGT	AAAAGGTAAG	TATATGACCG	CGACCTTGGT	CAGTGCAAAG	400
	ACCGGTGAAA	TCCTCGCTAC	CACCCAACGA	CCTACCTTTA	ATGCAGATAC	450
	TAAAGAAGGA	ATCACTGAGG	ACTTTGTTTG	GCGTGATATT	CTTTATCAAA	500
	GTAACATATGA	ACCAGGATCA	GCCATGAAGG	TTATGACGTT	AGCTTCTTCT	550
30	ATTGATAATA	ATACCTTCCC	AAGTGGAGAA	TACTTCAATA	GCAGTGAATT	600
	CAAAATAGCG	GATGCGACGA	CTCGAGATTG	GGATGTTAAT	GATGGTTTGA	650
	CTACTGGTGG	GATGATGACT	TTCTTACAAG	GTTTCGCTCA	CTCCAGTAAT	700
	GTTGGAATGA	GTCTACTTGA	ACAAAAAATG	GGAGATGCTA	CTTGTTTGA	750
	TTATCTAAAA	CGCTTTAAAT	TTGGGGTTCC	AACTCGCTTT	GGCTTGACAG	800
35	ATGAATACGC	TGGTCAACTT	CCAGCTGATA	ATATTGTTAG	TATTGCTCAA	850
	AGCTCATTTG	GGCAAGGAAT	TTCAGTGACA	CAAACACAAA	TGCTTCGTGC	900
	CTTTACAGCT	ATTGCTAATG	ATGGAGTTAT	GCTGGAGCCA	AAATTTATAA	950
	GTGCTATTTA	TGATACTAAC	AATCAGTCTG	TACGTAAGTC	ACAAAAAGAA	1000
	ATAGTAGGAA	ATCCTGTTTC	CAAAGAGGCA	GCAAGCACAA	CTCGAAATCA	1050
40	CATGATCTTA	GTTGGGACGG	ACCCTCTATA	TGGAACATATG	TATAATCACT	1100
	ACACAGGAAA	GCCAATTATA	ACAGTTCCTG	GACAAAATGT	AGCAGTTAAA	1150
	TCCGGTACGG	CTCAAATCGC	TGATGAGAAA	AATGGAGGAT	ACTTGGTTGG	1200
	TTCTACCAAT	TATATTTTCT	CAGTTGTGAC	TATGAATCCT	GCTGAAAATC	1250
	CTGATTTTAT	CTTGTATGTA	ACGGTTCAAC	AGCCTGAGCA	TTATTCAGGT	1300
45	ATCCAGTTGG	GAGAATTTGC	CACCCCAATC	TTGGAGCGGG	CTTCAGCTAT	1350
	GAAAGAATCT	CTCAATCTTC	AATCTCCAGC	CAAAAATTTA	GATAAAGTTA	1400
	CGACAGAATC	TTCTTATGCA	ATGCCTAGCA	TCAAGGATAT	TTCACCTGGT	1450
	GAGTTGGCGG	AAGCCTTACG	CCGAAATATT	GTGCAACCAA	TCGTTGTAGG	1500
	TACTGGAACA	AAGATTAAAG	AGACTTCTGT	AGAAGAAGGG	ACCAATCTTG	1550
50	CACCAAACCA	ACAAGTTCTC	CTTTTATCGG	ATAAGGTAGA	AGAAATTCCA	1600
	GACATGTATG	GCTGGAAAAA	AGAGACTGCC	GAGACCTTTG	CTAAATGGTT	1650
	GGATATTGAA	CTGGAATTTG	AAGGTTTCAGG	TTCCGTTGTT	CAGAAG	1696

## 2) INFORMATION FOR SEQ ID NO: 1042

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1689 bases  
 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-09

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1042

	AAAAAAGAGT	TGGAAACTGC	AGAGGTCAAG	GGGATTGATT	TTACAACCAG	50
	TCCCAATCGT	AGTTACCCAA	ACGGACAATT	TGCTTCTAGT	TTTATCGGAC	100
	TAGCTCAGCT	CCATGAAAAT	GAAGATGGCA	GCAAGAGCTT	GCTGGGAACC	150
15	TCTGGAATGG	AGAGTTCCTT	GAACAGTATT	CTTGCAGGGA	CAGACGGCAT	200
	TATTACCTAT	GAAAAGGATC	GTCTGGGAAA	TATTGTCCCC	GGAACGGAAC	250
	AAGTTTCCCA	ACAAACTGTA	GATGGCAAGG	ATGTTTATAC	GACTATTTCC	300
	AGCACCTTTC	AGTCCTTCAT	GGAAACTCAG	ATGGATGCCT	TTCTAGAAAA	350
	AGTAAAAGGT	AAGTATATGA	CCGCGACCTT	GGTCAGTGCA	AAGACCGGTG	400
20	AAATCCTCGC	TACCACCCAA	CGACCTACCT	TTAATGCAGA	TACTAAAGAA	450
	GGAATCACTG	AGGACTTTGT	TTGGCGTGAT	ATTCTTTATC	AAAGTAACTA	500
	TGAACCAGGA	TCAGCCATGA	AGGTTATGAC	GTTAGCTTCT	TCTATTGATA	550
	ATAATACCTT	CCCAAGTGGA	GAATACTTCA	ATAGCAGTGA	ATTCAAATA	600
	GCGGATGCGA	CGACTCGAGA	TTGGGATGTT	AATGATGGTT	TGACTACTGG	650
25	TGGGATGATG	ACTTTCTTAC	AAGGTTTCGC	TCACTCCAGT	AATGTTGGAA	700
	TGAGTCTACT	TGAACAAAAA	ATGGGAGATG	CTACTTGGTT	GGATTATCTA	750
	AAACGCTTTA	AATTTGGGGT	TCCAACCTCG	TTTGGCTTGA	CAGATGAATA	800
	CGCTGGTCAA	CTTCCAGCTG	ATAATATTGT	TAGTATTGCT	CAAAGCTCAT	850
	TTGGGCAAGG	AATTTTCAGT	ACACAAACAC	AAATGCTTCG	TGCCCTTTACA	900
30	GCTATTGCTA	ATGATGGAGT	TATGCTGGAG	CCAAAAATTA	TAAGTGCTAT	950
	TTATGATACT	AACAATCAGT	CTGTACGTAA	GTCACAAAAA	GAAATAGTAG	1000
	GAAATCCTGT	TTCCAAAGAG	GCAGCAAGCA	CAACTCGAAA	TCACATGATC	1050
	TTAGTTGGGA	CGGACCCTCT	ATATGGAAC	ATGTATAATC	ACTACACAGG	1100
	AAAGCCAATT	ATAACAGTTC	CTGGACAAAA	TGTAGCAGTT	AAATCCGGTA	1150
35	CGGCTCAAAT	CGCTGATGAG	AAAAATGGAG	GATACTTGGT	TGGTTCCTACC	1200
	AATTATATTT	TCTCAGTTGT	GACTATGAAT	CCTGCTGAAA	ATCCTGATTT	1250
	TATCTTGAT	GTAACGGTTC	AACAGCCTGA	GCATTATTCA	GGTATCCAGT	1300
	TGGGAGAATT	TGCCACCCCA	ATCTTGGAGC	GGGCTTCAGC	TATGAAAGAC	1350
	TCTCTCAATC	TTCAAACAAC	AGCTAAAGCT	TTGGAGCAAG	TAAGTCAACA	1400
40	AAGTCCCTTAT	CCTATGCCTA	GTGTCAAGGA	TATTTACCTT	GGTGATTTAG	1450
	CAGAAGAATT	GCGTCGCAAT	CTTGTAACAAC	CCATCGTTGT	GGGAACAGGA	1500
	ACGAAGATTA	AAAACAGTTC	TGCTGAAGAA	GGGAAGAATC	TTGCCCCGAA	1550
	TCAGCAAGTC	CTTATCTTAT	CTGATAAAGT	AGAGGAAGTT	CCAGATATGT	1600
	ATGGTTGGAC	AAAGGAGACT	GCTGAGACCC	TTGCTAAGTG	GCTCAATATA	1650
45	GAACTTGAAT	TTCAAGGTTC	GGGTTCTACT	GTGCAGAAG		1689

2) INFORMATION FOR SEQ ID NO: 1043

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1690 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

## (B) STRAIN: StrR-10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1043

5	CAAAAAAGAG	TTGGAAACTG	CAGAGGTC	GGGGATTGAT	TTTACAACCA	50
	GTCCTAATCG	TAGTTACCCA	AACGGACAAT	TTGCTTCTAG	TTTTATCGGT	100
	CTAGCTCAAC	TCCATGAAAA	TGAAGATGGA	AGCAAGAGTT	TGCTGGGAAC	150
	TTCTGGAATG	GAGAGTTCCT	TGAACAGTAT	TCTTGCAGGG	AAAGACGGTA	200
	TTATTACTTA	TGAAAAGGAT	CGTCTGGGTA	ATATTGTCCC	TGGAACAGAA	250
10	CAAGTTTCCC	AACAAACGGT	AGATGGCAAG	GATGTTTATA	CAACATTGTC	300
	TAGTCCGCTC	CAATCTTTCA	TGGAAACTCA	GATGGATGCC	TTTCTAGAAA	350
	AAGTAAAAGG	TAAGTATATG	ACCGCGACCT	TGGTCAGTGC	AAAGACCGGT	400
	GAAATCCTCG	CTACCACCCA	ACGACCTACC	TTTAATGCAG	ATACTAAAGA	450
	AGGAATCACT	GAGGACTTTG	TTTGGCGTGA	TATTCTTTAT	CAAAGTAACT	500
15	ATGAACCAGG	ATCAGCCATG	AAGGTTATGA	CGTTAGCTTC	TTCTATTGAT	550
	AATAATACCT	TCCCAAGTGG	AGAATACTTC	AATAGCAGTG	AATTCAAAAT	600
	AGCGGATGCG	ACGACTCGAG	ATTGGGATGT	TAATGATGGT	TTGACTACTG	650
	GTGGGATGAT	GACTTTCCTA	CAAGGTTTCG	CTCACTCCAG	TAATGTTGGA	700
	ATGAGTCTAC	TTGAACAAAA	AATGGGAGAT	GCTACTTGGT	TGGATTATCT	750
20	AAAACGCTTT	AAATTGGGGG	TTCCAACCTG	CTTTGGCTTG	ACAGATGAAT	800
	ACGCTGGTCA	ACTTCCAGCT	GATAATATTG	TTAGTATTGC	TCAAAGCTCA	850
	TTTGGGCAAG	GAATTTCAGT	GACACAAACA	CAAATGCTTC	GTGCCTTTAC	900
	AGCTATTGCT	AATGATGGAG	TTATGCTGGA	GCCAAAATTT	ATAAGTGCTA	950
	TTTATGATAC	TAACAATCAG	TCTGTACGTA	AGTCACAAAA	AGAAATAGTA	1000
25	GGAAATCCTG	TTTCCAAAGA	GGCAGCAAGC	ACAACCTCGAA	ATCACATGAT	1050
	CTTAGTTGGG	ACGGACCCTC	TATATGGAAC	TATGTATAAT	CACTACACAG	1100
	GAAAGCCAAAT	TATAACAGTT	CCTGGACAAA	ATGTAGCAGT	TAAATCCGGT	1150
	ACGGCTCAAA	TCGCTGATGA	GAAAAATGGA	GGATACTTGG	TTGGTCTCTAC	1200
	CAATTATATT	TTCTCAGTTG	TGACTATGAA	TCCTGCTGAA	AATCCTGATT	1250
30	TTATCTTGTA	TGTAACGGTT	CAACAGCCTG	AGCATTATTC	AGGTATCCAG	1300
	TTGGGAGAAT	TTGCCACCCC	AATCTTGGAG	CGGGCTTCAG	CTATGAAAGA	1350
	ATCTCTCAAT	CTTCAATCTC	CAGCCAAAAA	TTTAGATAAA	GTTACGACAG	1400
	AATCTTCTTA	TGCAATGCCT	AGCATCAAGG	ATATTTCAAC	TGGTGAGTTG	1450
	GCGGAAGCCT	TACGCCGAAA	TATTGTGCAA	CCAATCGTTG	TAGGTACTGG	1500
35	AACAAAGATT	AAAGAGACTT	CTGTAGAAGA	AGGGACCAAT	CTTGCACCAA	1550
	ACCAACAAGT	TCTCCTTTTA	TCGGATAAGG	TAGAAGAAAT	TCCAGACATG	1600
	TATGGCTGGA	AAAAAGAGAC	TGCCGAGACC	TTTGCTAAAT	GGTTGGATAT	1650
	TGAACCTGGA	TTTGAAGGTT	CAGGTTCCGT	TGTTCAGAAG		1690

40

## 2) INFORMATION FOR SEQ ID NO: 1044

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 1668 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 50 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-11

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1044

	AGTTGGAAAC	TGCAGAGGTC	AAGGGGATTG	ATTTTACAAC	CAGTCCTAAT	50
	CGTAGTTACC	CAAACGGACA	ATTTGCTTCT	AGTTTATATCG	GACTAGCTCA	100
60	GCTCCATGAA	AATGAAGATG	GCAGCAAGAG	CTTGCTGGGA	ACCTCTGGGA	150

	TGGAGAGTTC	TTTAAATAGA	ATTCTTGCAG	GGACAGACGG	CATTATTACC	200
	TATGAAAAGG	ATCGTCTGGG	AAATATTGTC	CCCGGAACGG	AACAAGTTTC	250
	CCAACAAACG	GTAGATGGCA	AGGATGTTTA	TACGACTATT	TCCAGCACCC	300
	TTCAGTCCTT	CATGGAGACC	CAGATGAATG	CCTTTCAAGA	AAAAGTAAAA	350
5	GGCAAGTATA	TGACGGCTAC	CTTGGTCAGT	GCTAAAACAG	GGGAAATTCT	400
	TGCAACAACG	CAACGGCCGA	CCTTCGATGC	TGATACTAAG	GAAGGACTTA	450
	CCAAGGACTT	TGTTTGGCGT	GATATCCTCT	ATCAAAGTAA	CTATGAGCCA	500
	GGGTCAACCA	TGAAGGTCAT	GACGCTTGCT	GCTGCTATTG	ATAATAACAC	550
	TTTCCCAGGA	GGAGAAGTTT	TCAATAGTAG	TGAATTAAAA	ATAGCGGATG	600
10	CGACAATTCT	AGATTGGGAT	GTTAATGATG	GTTTGACGAC	TGGTGCGGATG	650
	ATGACTTTCT	TACAAGGTTT	CGCTCACTCC	AGTAATGTTG	GAATGAGTCT	700
	ACTTGAACAA	AAAATGGGAG	ATGCTACTTG	GTTGGATTAT	CTAAACCGCT	750
	TTAAGTTTGG	GGTGCCGACG	CGTTTTGGTC	TGACTGATGA	GTATTCAGGT	800
	CAATTGCCTG	CAGATAATAT	TGTTAATATT	GCCATGAGTG	CATTTGGTCA	850
15	GGGGATTTC	GTGACCCAGA	CGCAAATGAT	TCGTGCCTTT	ACAGCTATTG	900
	CTAATGATGG	TGTTATGCTG	GAGCCTAAAT	TTATTAGTGC	CATTTATGAT	950
	CCAAATGATC	AAACTGCTCG	GAAATCTCAA	AAAGAAATTG	TGGGAAATCC	1000
	TGTTTCTAAA	GATGCAGCTA	GTCTAACTCG	GACTAACATG	ATTTTGGTAG	1050
	GGACGGATCC	GGTTTATGGA	ACCATGTATA	ACCACAGCAC	AGGTAAGCCA	1100
20	ACTGTAACCTG	TTCTTGGGCA	AAATGTAGCC	CTCAAGTCTG	GTACGGCTCA	1150
	GATTGCTGAC	GAGAAAAATG	GTGGTTATCT	AGTCGGGTTA	ACCAACTATA	1200
	TTTTCTCGCG	TGTATCGATG	AATCCGGCTG	AAAATCCTGA	TTTTATCTTG	1250
	TATGTAACGG	TACAGCAACC	TGAACATTAT	TCAGGTATCC	AGTTGGGAGA	1300
	ATTTGCCACC	CCAATCTTGG	AGCGGGCTTC	AGCTATGAAA	GAATCTCTCA	1350
25	ATCTTCAATC	TCCAGCCAAG	AATTTAGATA	AAGTTACGAC	AGAATCTTCT	1400
	TATGCAATGC	CTAGCATCAA	GGATATTTCA	CCTGGTGAGT	TGGCGGAAGC	1450
	CTTACGCCGA	AATATTGTGC	AACCAATCGT	TGTAGGTACT	GGAACAAAGA	1500
	TTAAAGAGAC	TTCTGTAGAA	GAAGGGACCA	ATCTTGCACC	AAACCAACAA	1550
	GTTCTCCTTT	TATCGGATAA	GGTAGAAGAA	ATTCCAGACA	TGTATGGCTG	1600
30	GAAAAAAGAG	ACTGCTGAAA	CCTTTGCTAA	ATGGTTGAT	ATTGAGCTGG	1650
	AATTTGAAGG	GTCAGGTT				1668

## 35 2) INFORMATION FOR SEQ ID NO: 1045

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 1680 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 45 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: StrR-12

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1045

50	TTGGAAACTG	CAGAGGTCAA	GGGGATTGAT	TTTACAACCA	GTCCCAATCG	50
	TAGTTACCCA	AATGGACAAT	TTGCTTCTAG	TTTTATCGGT	CTAGCTCAGC	100
	TCCATGAAAA	TGAAGATGGA	AGCAAGAGTT	TGCTGGGAAC	CTCTGGAATG	150
	GAGAGTTCCT	TGAACAGTAT	TCTTGCAGGG	ACAGACGGCA	TTATTACCTA	200
55	TGAAAAGGAT	CGTCTGGGCA	ATATTGTACC	CGGAACAGAA	CAAGTTTCCC	250
	AACAAACGGT	GGATGGCAAG	GATGTTTACA	CAACCATTTC	CAGCCCCCTC	300
	CAGTCCTTCA	TGGAAACTCA	GATGGATGCC	TTTCTAGAAA	AAGTAAAAGG	350
	TAAGTATATG	ACCGCGACCT	TGGTCAGTGC	AAAGACCGGT	GAAATCCTCG	400
	CTACCACCCA	ACGACCTACC	TTTAATGCAG	ATACTAAAGA	AGGAATCACT	450
60	GAGGACTTTG	TTTGGCGTGA	TATTCTTTAT	CAAAGTAACT	ATGAACCAGG	500

	ATCAGCCATG	AAGGTTATGA	CGTTAGCTTC	TTCTATTGAT	AATAATACCT	550
	TCCCAAGTGG	AGAATACCTC	AATAGCAGTG	AATTCAAAAT	AGCGGATGCG	600
	ACGACTCGAG	ATTGGGATGT	TAATGATGGT	TTGACTACTG	GTGGGATGAT	650
	GACTTTCTTA	CAAGGTTTCG	CTCACTCCAG	TAATGTTGGA	ATGAGTCTAC	700
5	TTGAACAAAA	AATGGGAGAT	GCTACTTGGT	TGGATTATCT	AAAACGCTTT	750
	AAATTTGGGG	TTCCAACCTC	CTTTGGCTTG	ACAGATGAAT	ACGCTGGTCA	800
	ACTTCCAGCT	GATAATATTG	TTAGTATTGC	TCAAAGCTCA	TTTGGGCAAG	850
	GAATTTTCAGT	GACACAAACA	CAAATGCTTC	GTGCCTTTAC	AGCTATTGCT	900
	AATGATGGAG	TTATGCTGGA	GCCAAAATTT	ATAAGTGCTA	TTTATGATAC	950
10	TAACAATCAG	TCTGTACGTA	AGTCACAAAA	AGAAATAGTA	GGAAATCCTG	1000
	TTTCCAAAGA	GGCAGCAAGC	ACAACTCGAA	ATCACATGAT	CTTAGTTGGG	1050
	ACGGACCCTC	TATATGGAAC	TATGTATAAT	CACACACAG	GAAAGCCAAT	1100
	TATAACAGTT	CCTGGACAAA	ATGTAGCAGT	TAAATCCGGT	GCGGCTCAAA	1150
	TCGCTGATGA	GAAAAATGGA	GGATACTTGG	TTGGTTCTAC	CAATTATATT	1200
15	TTCTCAGTTG	TGACTATGAA	TCCTGCTGAA	AATCCTGATT	TTATCTTGTA	1250
	TGTAACGGTT	CAACAGCCTG	AGCATTATTC	AGGTATCCAG	TTGGGAGAAT	1300
	TTGCCACCCC	AATCTTGGAG	CGGGCTTCAG	CTATGAAAGA	ATCTCTCAAT	1350
	CTTCAATCTC	CAGCCAAAAA	TTTAGATAAA	GTTACGACAG	AATCTTCTTA	1400
	TGCAATGCCT	AGCATCAAGG	ATATTTCCAC	TGGTGAGTTG	GCGGAAGCCT	1450
20	TACGCCGAAA	TATTGTGCAA	CCAATCGTTG	TAGGTACTGG	AACAAAGATT	1500
	AAAGAGACTT	CTGTAGAAGA	AGGGACCAAT	CTTGCACCAA	ACCAACAAGT	1550
	TCTCCTTTTA	TCGGATAAGG	TAGAAGAAAT	TCCAGACATG	TATGGCTGGA	1600
	AAAAAGAGAC	TGCCGAGACC	TTTGCTAAAT	GGTTGGATAT	TGAACTGGAA	1650
	TTTGAAGGTT	CAGGTTCCGT	TGTTTCAGAAG			1680

## 2) INFORMATION FOR SEQ ID NO: 1046

30 (i) SEQUENCE CHARACTERISTICS: ---

(A) LENGTH: 1689 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*

40 (B) STRAIN: StrR-13

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1046

	AAAAAAGAGT	TGGAAACTGC	AGAGGTCAAG	GGGATTGATT	TTACAACCAG	50
45	TCCTAATCGT	AGTTACCCAA	ACGGACAATT	TGCTTCTAGT	TTTATCGGTC	100
	TAGCTCAACT	CCATGAAAAT	GAAGATGGAA	GCAAGAGTTT	GCTGGGAACT	150
	TCTGGAATGG	AGAGTTCCTT	GAACAGTATT	CTTGCAGGGA	AAGACGGTAT	200
	TATTACTTAT	GAAAAGGATC	GTCTGGGTAA	TATTGTCCCT	GGAACAGAAC	250
	AAGTTTCCCA	ACAAACGGTA	GATGGCAAGG	ATGTTTATAC	AACATTGTCT	300
50	AGTCCGCTCC	AATCTTTTCAT	GGAAACTCAG	ATGGATGCCT	TTCTAGAAAA	350
	AGTAAAAGGT	AAGTATATGA	CCGCGACCTT	GGTCAGTGCA	AAGACCGGTG	400
	AAATCCTCGC	TACCACCCAA	CGACCTACCT	TTAATGCAGA	TACTAAAGAA	450
	GGAATCACTG	AGGACTTTGT	TTGGCGTGAT	ATTCTTTATC	AAAGTAACTA	500
	TGAACCAGGA	TCAGCCATGA	AGGTTATGAC	GTTAGCTTCT	TCTATTGATA	550
55	ATAATACCTT	CCCAAGTGGA	GAATACTTCA	ATAGCAGTGA	ATTCAAAATA	600
	GCGGATGCGA	CGACTCGAGA	TTGGGATGTT	AATGATGGTT	TGACTACTGG	650
	TGGGATGATG	ACTTTCTTAC	AAGGTTTCGC	TCACTCCAGT	AATGTTGGAA	700
	TGAGTCTACT	TGAACAAAAA	ATGGGAGATG	CTACTTGGTT	GGATTATCTA	750
	AAACGCTTTA	AATTTGGGGT	TCCAACTCGC	TTTGGCTTGA	CAGATGAATA	800
60	CGCTGGTCAA	CTTCCAGCTG	ATAATATTGT	TAGTATTGCT	CAAAGCTCAT	850



	TTGGGCAAGG	AATTTTCAGTG	ACACAAACAC	AAATGCTTCG	TGCCTTTACA	900
	GCTATTGCTA	ATGATGGAGT	TATGCTGGAG	CCAAAATTTA	TAAAGTGCTAT	950
	TTATGATACT	AACAATCAGT	CTGTACGTAA	GTCACAAAAA	GAAATAGTAG	1000
	GAAATCCTGT	TTCCAAAGAG	GCAGCAAGCA	CAACTCGAAA	TCACATGATC	1050
5	TTAGTTGGGA	CGGACCTCT	ATATGGAAC	ATGTATAATC	ACTACACAGG	1100
	AAAGCCAATT	ATAACAGTTC	CTGGACAAAA	TGTAGCAGTT	AAATCCGGTA	1150
	CGGCTCAAA	CGCTGATGAG	AAAAATGGAG	GATACTTGGT	TGGTTCTACC	1200
	AATTATATTT	TCTCAGTTGT	GACTATGAAT	CCTGCTGAAA	ATCCTGATTT	1250
	TATCTTGTAT	GTAACGGTTC	AACAGCCTGA	GCATTATTCA	GGTATCCAGT	1300
10	TGGGAGAATT	TGCCACCCCA	ATCTTGGAGC	GGGCTTCAGC	TATGAAAGAA	1350
	TCTCTCAATC	TTCAATCTCC	AGCCAAAAAT	TTAGATAAAG	TTACGACAGA	1400
	ATCTTCTTAT	GCAATGCCTA	GCATCAAGGA	TATTTACCT	GGTGAGTTGG	1450
	CGGAAGCCTT	ACGCCGAAAT	ATTGTGCAAC	CAATCGTTGT	AGGTACTGGA	1500
	ACAAAGATTA	AAGAGACTTC	TGTAGAAGAA	GGGACCAATC	TTGCACCAAA	1550
15	CCAACAAGTT	CTCCTTTTAT	CGGATAAGGT	AGAAGAAATT	CCAGACATGT	1600
	ATGGCTGGAA	AAAAGAGACT	GCCGAGACCT	TTGCTAAATG	GTTGGATATT	1650
	GAACTGGAAT	TTGAAGGTTT	AGGTTCCGTT	GTTTACAAG		1689

20

## 2) INFORMATION FOR SEQ ID NO: 1047

## (i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 1690 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

30

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*
- (B) STRAIN: StrR-14

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1047

	CAAAAAAGAA	TTGGAAGCTG	CAGAGGTC	GGGGATTGAT	TTTACAACCA	50
	GTCCCAACCG	TAGTTACCCA	AACGGACAAT	TTGCTTCTAG	TTTTATCGGA	100
	CTAGCTCAGC	TCCATGAAAA	TGAAGATGGC	AGCAAGAGCT	TGCTGGGAAC	150
40	TTCTGGAATG	GAGAGTTCCT	TGAACAGTAT	TCTTGCAGGG	ACAGACGGCA	200
	TTATTACCTA	TGAAAAGGAT	CGTCTGGGTA	ATATTGTACC	CGGAACAGAA	250
	CAAGTTTCCC	AACAAACGGT	AGATGGCAAG	GATGTTTACA	CAACCATTTC	300
	CAGCCCCCTC	CAGTCCTTTA	TGGAAACCCA	GATGGATGCT	TTTCAAGAGA	350
	AGGTAAAAGG	AAAGTACACG	ACAGCGACTT	TGGTCAGTGC	TAAAACAGGG	400
45	GAAATTCTGG	CAACAACGCA	ACGACCGACC	TTTGATGCAG	ATACAAAAGA	450
	AGGCATTACA	GAGGACTTTG	TTTGGCGTGA	TATCCTTTAC	CAAAGTAACT	500
	ATGAGCCAGG	TTCCCCTATG	AAAGTGATGA	TGTTGGCTGC	TGCTATTGAT	550
	AATAATACCT	TTCCAGGGGG	AGAAGTCTTC	AATAGTAGTG	AGTTAAAAAT	600
	TGCAGATGTC	ACGATTCGAG	ATTGGGACGT	CAATGAAGGA	TTGACTGGTG	650
50	GCAGAATGAT	GACCTTTTCT	CAAGGGTTTC	CTCACTCAAG	TAACGTTGGG	700
	ATGACGCTTC	TTGAGCAAAA	GATGGGAGAT	GCTACATGGC	TTGATTATCT	750
	AAATCGCTTT	AAATTTGGTG	TTCCGACTCG	TTTTGGCTTG	ACGGATGAAT	800
	ATGCAGGCCA	ACTTCCAGCT	GACAATATCG	TAAATATTGC	TCAGAGTTCA	850
	TTTGGACAAG	GGATTTCAGC	GACCCAGACG	CAAATGATTC	GTGCCTTCAC	900
55	GGCTATTGCC	AACGATGGAG	TCATGTTAGA	ACCTAAATTT	ATCAGTGCCA	950
	TTTATGATCC	AAATGATCAA	ACTGCTCGGA	AATCACAAAA	AGAAGTTGTG	1000
	GGAAATCCTG	TGTCTAAAGA	TGCAGCGAGC	TTGACGCGAA	CGCATATGGT	1050
	TTTAGTCGGT	ACCGATCCAG	TATATGGAAC	TATGTATAAT	CATAAGACAG	1100
	GGAAACCAAC	TGTAACGTGT	CCTGGGCAAA	ATGTAGCCCT	CAAGTCTGGT	1150
60	ACGGCTCAGA	TTGCCGATGA	GAAAAATGGA	GTTTACTTAG	TTGGTACGAC	1200

	CAATTACATT	TTTTCGGCTG	TATCGATGAA	CCCTGCTGAA	AATCCTGATT	1250
	TTATTCTCTA	TGTGACGGTT	CAACAGCCTG	AGCATTATTC	AGGTATTCAG	1300
	TTGGGGGAAT	TTGCCAATCC	TATCTTGGAA	AGGGCAGTGG	CTATGAAAGA	1350
5	TTCCCTTAAC	CTCCAATCTA	CCGCTAAAAC	GTTAAATCAG	GTAACCAATC	1400
	AAAGCGCTTA	TGCCATGCCT	AGCATCAAGG	ACATTTTACC	TGGCGATTTG	1450
	GCGGAAGCCT	TACGTGCGAA	TATTGTGCAA	CCAATCGTTG	TAGGAACAGG	1500
	AACAAAAATT	AAAGAATCAT	CTGTAGAAGA	AGGGACGGAT	CTTGCACCTA	1550
	ACCAGCAAGT	TCTTCTCTTA	TCTGATAAAG	CAGAGGAAGT	TCCAGATATG	1600
	TATGGTTGGA	CAAAAGAGAC	TGCTGAGACC	TTTGCTAAGT	GGCTCAATAT	1650
10	AGAACTTGAA	TTTGAAGGTT	CGGGCTCTAC	TGTGCAGAAG		1690

## 2) INFORMATION FOR SEQ ID NO: 1048

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1682 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1048

30	AAAAGAGTTG	GAAACTGCAG	AGGTCAAGGG	GATTGATTTT	ACAACCAGTC	50
	CCAATCGTAG	TTACCCAAAC	GGACAATTTG	CTTCTAGCTT	TATTGGCTTA	100
	GCCCAACTTC	ATGAAAATGA	GGATGGTAGT	AAGAGTTTGT	TAGGGACTTC	150
	TGGTTTGGAG	AGTTCTTTAA	ATACCATTCT	TGCTGGGACA	GACGGTATTA	200
	TTACCTATGA	AAAAGACCGT	GTAGGAAATA	TCGTACCAGG	TACAGAACTG	250
35	GTATCGCAAC	AAACTGTGGA	TGGCAAGGAT	GTTTATACAA	CATTGTCTAG	300
	TCCGCTACAA	TCTTTCATGG	AAACTCAGAT	GGATGCCTTT	CTAGAAAAAG	350
	TAAAAGGTAA	GTATATGACC	GCGACCTTGG	TCAGTGCAAA	GACCGGTGAA	400
	ATCCTCGCTA	CCACCCAACG	ACCTACCTTT	AATGCAGATA	CTAAAGAAGG	450
	AATCACTGAG	GACTTTGTTT	GGCGTGATAT	TCTTTATCAA	AGTAACTATG	500
40	AACCAGGATC	AGCCATGAAG	GTTATGACGT	TAGCTTCTTC	TATTGATAAT	550
	AATACCTTCC	CAAGTGGAGA	ATACTTCAAT	AGCAGTGAAT	TCAAAATAGC	600
	GGATGCGACG	ACTCGAGATT	GGGATGTTAA	TGAAGGTTTG	ACTACTGGTG	650
	GGATGATGAC	TTTCTTACAA	GTTTTCGCTC	ACTCCAGTAA	TGTTGGAATG	700
	AGTCTACTTG	AACAAAAAAT	GGGAGATGCT	ACTTGGTTGG	ATTATCTAAA	750
45	ACGCTTTAAA	TTTGGGGTTC	CAACTCGCTT	TGGCTTGACA	GATGAATACG	800
	CTGGTCAACT	TCCAGCTGAT	AATATTGTTA	GTATTGCTCA	AAGCTCATTT	850
	GGGCAAGGAA	TTTCAGTGAC	ACAAACACAA	ATGCTTCGTG	CCTTTACAGC	900
	TATTGCTAAT	GATGGAGTTA	TGCTGGAGCC	AAAATTTATA	AGTGCTATTT	950
	ATGATACTAA	CAATCAGTCT	GTACGTAAGT	CACAAAAAGA	AATAGTAGGA	1000
50	AATCCTGTTT	CCAAAGAGGC	AGCAAGCACA	ACTCGAAATC	ACATGATCTT	1050
	AGTTGGGACG	GACCCCTCTAT	ATGGAACATAT	GTATAATCAC	TACACAGGAA	1100
	AGCCAATTAT	AACAGTTCCT	GGACAAAATG	TAGCAGTTAA	ATCCGGTACG	1150
	GCTCAAATCG	CTGATGAGAA	AAATGGAGGA	TACTTGGTTG	GTTCTACCAA	1200
	TTATATTTTC	TCAGTTGTGA	CTATGAATCC	TGCTGAAAAT	CCTGATTTTA	1250
55	TCTTGATATG	AACGGTTCAA	CAGCCTGAGC	ATTATTCAGG	TATCCAGTTG	1300
	GGAGAATTTG	CCACCCCAAT	CTTGGAGCGG	GCTTCAGCTA	TGAAAGAATC	1350
	TCTCAATCTT	CAATCTCCAG	CCAAAAATTT	AGATAAAGTT	ACGACAGAAAT	1400
	CTTCTTATGC	AATGCCTAGC	ATCAAGGATA	TTTCACCTGG	TGAGTTGGCG	1450
	GAAGCCTTAC	GCCGAAATAT	TGTGCAACCA	ATCGTTGTAG	GTAAGTGAAC	1500
60	AAAGATTAAA	GAGACTTCTG	TAGAAGAAGG	GACCAATCTT	GCACCAAACC	1550

AACAAGTTCT	CCTTTTATCG	GATAAGGTAG	AAGAAATTCC	AGACATGTAT	1600
GGCTGGAAAA	AAGAGACTGC	TGAAACCTTT	GCTAAATGGT	TGGATATTGA	1650
GTTGGAATTT	GAAGGTCAG	GTTCCGTCGT	TC		1682

5

## 2) INFORMATION FOR SEQ ID NO: 1049

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 1241 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 15 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: R690

20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1049

GCAGCGTTGC	GTGATACCGT	TGAAAAAACC	ATTAAAAACT	GTTTGGATTT	50
TGAAAGGAGA	CAGGAGCATG	AATAGAATAA	AAGTTGCAAT	ACTGTTTGGG	100
25 GGTGCTCAG	AGGAGCATGA	CGTATCGGTA	AAATCTGCAA	TAGAGATAGC	150
CGCTAACATT	AATAAAGAAA	AATACGAGCC	GTTATACATT	GGAATTACGA	200
AATCTGGTGT	ATGGAAAATG	TGCGAAAAAC	CTTGCGCGGA	ATGGGAAAAC	250
GACAATTGCT	ATTCAGCTGT	ACTCTCGCCG	GATAAAAAAA	TGCACGGATT	300
ACTTGTTAAA	AAGAACCATG	AATATGAAAT	CAACCATGTT	GATGTAGCAT	350
30 TTTCAGCTTT	GCATGGCAAG	TCAGGTGAAG	ATGGATCCAT	ACAAGGTCTG	400
TTTGAATGTG	CCGGTATCCC	TTTTGTAGGC	TGCGATATTC	AAAGCTCAGC	450
AATTTGTATG	GACAAATCGT	TGACATACAT	CGTTGCGAAA	AATGCTGGGA	500
TAGCTACTCC	CGCCTTTTGG	GTTATTAATA	AAGATGATAG	GCCGGTGGCA	550
GCTACGTTTA	CCTATCCTGT	TTTTGTTAAG	CCGGCGCGTT	CAGGCTCATC	600
35 CTTCGGTG TG	AAAAAAGTCA	ATAGCGCGGA	CGAATTGGAC	TACGCAATTG	650
AATCGGCAAG	ACAATATGAC	AGCAAAATCT	TAATTGAGCA	GGCTGTTTCG	700
GGCTGTGAGG	TCGGTTGTGC	GGTATTGGGA	AACAGTGCCG	CGTTAGCTGT	750
TGGCGAGGTG	GACCAATCA	GGCTGCAGTA	CGGAATCTTT	CGTATTCATC	800
AGGAAGTCGA	GCCGGAAAAA	GGCTCTGAAA	ACGCAGTTAT	AACCGTTCCC	850
40 GCAGACCTTT	CAGCAGAGGA	GCGAGGACGG	ATACAGGAAA	CGGCAAAAAA	900
AATATATAAA	GCGCTCGGCT	GTAGAGGTCT	AGCCCGTGTG	GATATGTTTT	950
TACAAGATAA	CGGCCGCATT	GTA CTGAACG	AAGTCAATAC	TCTGCCCGGT	1000
TTCACGTCAT	ACAGTCGTTA	TCCCCGTATG	ATGGCCGCTG	CAGGTATTGC	1050
ACTTCCCGAA	CTGATTGACC	GCTTGATCGT	ATTAGCGTTA	AAGGGGTGAT	1100
45 AAGCATGGAA	ATAGGATTTA	CTTTTTTTAGA	TGAAATAGTA	CACGGTGTTT	1150
GTTGGGACGC	TAAATATGCC	ACTTGGGATA	ATTTACCCGG	AAAACCGGTT	1200
GACGGTTATG	AAGTAAATCG	CATTGTAGGG	ACATACGAAT	T	1241

50

## 2) INFORMATION FOR SEQ ID NO: 1050

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1249 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

60

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus gallinarum*

(B) STRAIN: R691

## 5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1050

	TCACACCGCA	TACGGCCTAT	TATACCGAGC	AAGCGTTGCG	TGATACCGTT	50
	GAAAAAACCA	TTAAAACTG	TTTGGATTTT	GAAAGGAGAC	AGGAGCATGA	100
	ATAGAATAAA	AGTTGCAATA	CTGTTTGGGG	GTTGCTCAGA	GGAGCATGAC	150
10	GTATCGGTAA	AATCTGCAAT	AGAGATAGCC	GCTAACATTA	ATAAAGAAAA	200
	ATACGAGCCG	TTATACATTG	GAATTACGAA	ATCTGGTGTA	TGGAAAATGT	250
	GCGAAAAACC	TTGCGCGGAA	TGGGAAAACG	ACAATTGCTA	TTCAGCTGTA	300
	CTCTCGCCGG	ATAAAAAAAT	GCACGGATTA	CTTGTTAAAA	AGAACCATGA	350
	ATATGAAATC	AACCATGTTG	ATGTAGCATT	TTCAGCTTTG	CATGGCAAGT	400
15	CAGGTGAAGA	TGGATCCATA	CAAGGTCTGT	TTGAATTGTC	CGGTATCCCT	450
	TTTGTAGGCT	GCGATATTCA	AAGCTCAGCA	ATTTGTATGG	ACAAATCGTT	500
	GACATACATC	GTTGCGAAAA	ATGCTGGGAT	AGCTACTCCC	GCCTTTTGGG	550
	TTATTAATAA	AGATGATAGG	CCGGTGGCAG	CTACGTTTAC	CTATCCTGTT	600
	TTTGTTAAGC	CGGCGCGTTC	AGGCTCATCC	TTCGGTGTGA	AAAAAGTCAA	650
20	TAGCGCGGAC	GAATTGGACT	ACGCAATTGA	ATCGGCAAGA	CAATATGACA	700
	GCAAAATCTT	AATTGAGCAG	GCTGTTTCGG	GCTGTGAGGT	CGGTTGTGCG	750
	GTATTGGGAA	ACAGTGCCGC	GTTAGCTGTT	GGCGAGGTGG	ACCAAATCAG	800
	GCTGCAGTAC	GGAATCTTTC	GTATTCATCA	GGAAGTCGAG	CCGGAAAAAG	850
	GCTCTGAAAA	CGCAGTTATA	ACCGTTCCCG	CAGACCTTTC	AGCAGAGGAG	900
25	CGAGGACGGA	TACAGGAAAC	GGCAAAAAAA	ATATATAAAG	CGCTCGGCTG	950
	TAGAGGTCTA	GCCCGTGTGG	ATATGTTTTT	ACAAGATAAC	GGCCGCATTG	1000
	TACTGAACGA	AGTCAATACT	CTGCCCGGTT	TCACGTCATA	CAGTCGTTAT	1050
	CCCCGTATGA	TGGCCGCTGC	AGGTATTGCA	CTTCCCGAAC	TGATTGACCG	1100
	CTTGATCGTA	TTAGCGTTAA	AGGGGTGATA	AGCATGGAAA	TAGGATTTAC	1150
30	TTTTTTTAGAT	GAAATAGTAC	ACGGTGTTTCG	TTGGGACGCT	AAATATGCCA	1200
	CTTGGGATAA	TTTACCAGGA	AAACCGGTTG	ACGGTTATGA	AGTAAATCG	1249

## 35 2) INFORMATION FOR SEQ ID NO: 1051

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1272 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 45 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*

(B) STRAIN: R481

## 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1051

	TATACCGAGC	AAGCGTTGCG	TGATACCGTT	GAAAAAACCA	TTAAAACTG	50
	TTTGGATTTT	GAAAGGAGAC	AGGAGCATGA	ATAGAATAAA	AGTTGCAATA	100
	CTGTTTGGGG	GTTGCTCAGA	GGAGCATGAC	GTATCGGTAA	AATCTGCAAT	150
	AGAGATAGCC	GCTAACATTA	ATAAAGAAAA	ATACGAGCCG	TTATACATTG	200
55	GAATTACGAA	ATCTGGTGTA	TGGAAAATGT	GCGAAAAACC	TTGCGCGGAA	250
	TGGGAAAACG	ACAATTGCTA	TTCAGCTGTA	CTCTCGCCGG	ATAAAAAAAT	300
	GCACGGATTA	CTTGTTAAAA	AGAACCATGA	ATATGAAATC	AACCATGTTG	350
	ATGTAGCATT	TTCAGCTTTG	CATGGCAAGT	CAGGTGAAGA	TGGATCCATA	400
	CAAGGTCTGT	TTGAATTGTC	CGGTATCCCT	TTTGTAGGCT	GCGATATTCA	450
60	AAGCTCAGCA	ATTTGTATGG	ACAAATCGTT	GACATACATC	GTTGCGAAAA	500

```

ATGCTGGGAT AGCTACTCCC GCCTTTTGGG TTATTAATAA AGATGATAGG 550
CCGGTGGCAG CTACGTTTAC CTATCCTGTT TTTGTTAAGC CGGCGCGTTC 600
AGGCTCATCC TTCGGTGTGA AAAAAGTCAA TAGCGCGGAC GAATTGGACT 650
ACGCAATTGA ATCGGCAAGA CAATATGACA GCAAAATCTT AATTGAGCAG 700
5 GCTGTTTCGG GCTGTGAGGT CGGTTGTGCG GTATTGGGAA ACAGTGCCGC 750
GTTAGTTGTT GGCGAGGTGG ACCAAATCAG GCTGCAGTAC GGAATCTTTC 800
GTATTCATCA GGAAGTCGAG CCGGAAAAAG GCTCTGAAAA CGCAGTTATA 850
ACCGTTCCCG CAGACCTTTC AGCAGAGGAG CGAGGACGGA TACAGGAAAC 900
GGCAAAAAAA ATATATAAAG CGCTCGGCTG TAGAGGTCTA GCCCGTGTGG 950
10 ATATGTTTTT ACAAGATAAC GGCCGCATTG TACTGAACGA AGTCAATACT 1000
CTGCCCCGTT TCACGTCATA CAGTCGTTAT CCCCCTATGA TGGCCGCTGC 1050
AGGTATTGCA CTTCCCGAAC TGATTGACCG CTTGATCGTA TTAGCGTTAA 1100
AGGGGTGATA AGCATGGAAA TAGGATTTAC TTTTTTAGAT GAAATAGTAC 1150
ACGGTGTTCG TTGGGACGCT AAATATGCCA CTTGGGATAA TTTCACCGGA 1200
15 AAACCGGTTG ACGGTTATGA AGTAAATCGC ATTGTAGGGA CATACGAATT 1250
GGCTTGAATC GCTTTTTGAA GG 1272

```

20 2) INFORMATION FOR SEQ ID NO: 1052

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1237 bases  
 (B) TYPE: Nucleic acid  
 25 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: R492

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1052

```

35 TCCCCCGGCA TACGGCCTAT TATACCGAGC AAGCGTTGCG TGATACCGTT 50
GAAAAAACCA TTA AAAACTG TTTGGATTTT GAAAGGAGAC AGGAGCATGA 100
ATAGAATAAA AGTTGCAATA CTGTTTGGGG GTTGCTCAGA GGAGCATGAC 150
GTATCGGTAA AATCTGCAAT AGAGATAGCC GCTAACATTA ATAAAGAAAA 200
40 ATACGAGCCG TTATACATTG GAATTACGAA ATCTGGTGTA TGGAAAATGT 250
GCGAAAAACC TTGCGCGGAA TGGGAAAACG ACAATTGCTA TTCAGCTGTA 300
CTCTCGCCGG ATAAAAAAT GCACGGATTA CTTGTTAAAA AGAACCATGA 350
ATATGAAATC AACCATGTTG ATGTAGCATT TTCAGCTTTG CATGGCAAGT 400
CAGGTGAAGA TGGATCCATA CAAGGTCGTG TTGAATTGTC CGGTATCCCT 450
45 TTTGTAGGCT GCGATATTCA AAGCTCAGCA ATTTGTATGG ACAAATCGTT 500
GACATACATC GTTGCGAAAA ATGCTGGGAT AGCTACTCCC GCCTTTTGGG 550
TTATTAATAA AGATGATAGG CCGGTGGCAG CTACGTTTAC CTATCCTGTT 600
TTTGTTAAGC CGGCGCGTTC AGGCTCATCC TTCGGTGTGA AAAAAGTCAA 650
TAGCGCGGAC GAATTGGACT ACGCAATTGA ATCGGCAAGA CAATATGACA 700
50 GCAAAATCTT AATTGAGCAG GCTGTTTCGG GCTGTGAGGT CGGTTGTGCG 750
GTATTGGGAA ACAGTGCCGC GTTAGCTGTT GGCGAGGTGG ACCAAATCAG 800
GCTGCAGTAC GGAATCTTTC GTATTCATCA GGAAGTCGAG CCGGAAAAAG 850
GCTCTGAAAA CGCAGTTATA ACCGTTCCCG CAGACCTTTC AGCAGAGGAG 900
CGAGGACGGA TACAGGAAAC GGCAAAAAAA ATATATAAAG CGCTCGGCTG 950
55 TAGAGGTCTA GCCCGTGTGG ATATGTTTTT ACAAGATAAC GGCCGCATTG 1000
TACTGAACGA AGTCAATACT CTGCCCCGTT TCACGTCATA CAGTCGTTAT 1050
CCCCGTATGA TGGCCGCTGC AGGTATTGCA CTTCCCGAAC TGATTGACCG 1100
CTTGATCGTA TTAGCGTTAA AGGGGTGATA AGCATGGAAA TAGGATTTAC 1150
TTTTTTAGAT GAAATAGTAC ACGGTGTTCG TTGGGACGCT AAATATGCCA 1200
60 CTTGGGATAA TTTCACCGGA AAACCGGTTG ACGGTTA 1237

```

## 2) INFORMATION FOR SEQ ID NO: 1053

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1263 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: R581

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1053

20	CATACGGCCT	ATTATACCGA	GCAAGCGTTG	CGTGATACCG	TTGAAAAAAC	50
	CATTAAAAAC	TGTTTGGATT	TTGAAAGGAG	ACAGGAGCAT	GAATAGAATA	100
	AAAGTTGCAA	TACTGTTTGG	GGGTTGCTCA	GAGGAGCATG	ACGTATCGGT	150
	AAAATCTGCA	ATAGAGATAG	CCGCTAACAT	TAATAAAGAA	AAATACGAGC	200
	CGTTATACAT	TGGAATTACG	AAATCTGGTG	TATGGAAAAT	GTGCGAAAAA	250
25	CCTTGCGCGG	AATGGGAAAA	CGACAATTGC	TATTCAGCTG	TACTCTCGCC	300
	GGATAAAAAA	ATGCACGGAT	TACTTGTTAA	AAAGAACCAT	GAATATGAAA	350
	TCAACCATGT	TGATGTAGCA	TTTTTCAGCTT	TGCATGGCAA	GTCAGGTGAA	400
	GATGGATCCA	TACAAGGTCT	GTTTGAATTG	TCCGGTATCC	CTTTTGTAGG	450
	CTGCGATATT	CAAAGCTCAG	CAATTTGTAT	GGACAAATCG	TTGACATACA	500
30	TCGTTGCGAA	AAATGCTGGG	ATAGCTACTC	CCGCCTTTTG	GGTTATTAAT	550
	AAAGATGATA	GGCCGGTGCG	AGCTACGTTT	ACCTATCCTG	TTTTTGTAA	600
	GCCGCGCGGT	TCAGGCTCAT	CCTTCGGTGT	GAAAAAAGTC	AATAGCGCGG	650
	ACGAATTGGA	CTACGCAATT	GAATCGGCAA	GACAATATGA	CAGCAAAATC	700
	TTAATTGAGC	AGGCTGTTTC	GGGCTGTGAG	GTCGGTTGTG	CGGTATTGGG	750
35	AAACAGTGCC	GCGTTAGCTG	TTGGCGAGGT	GGACCAAATC	AGGCTGCAGT	800
	ACGGAATCTT	TCGTATTTCAT	CAGGAAGTCG	AGCCGGAAAA	AGGCTCTGAA	850
	AACGCAGTTA	TAACCGTTCC	CGCAGACCTT	TCAGCAGAGG	AGCGAGGACG	900
	GATACAGGAA	ACGGCAAAAA	AAATATATAA	AGCGCTCGGC	TGTAGAGGTC	950
	TAGCCCGTGT	GGATATGTTT	TTACAAGATA	ACGGCCGCAT	TGTACTGAAC	1000
40	GAAGTCAATA	CTCTGCCCCG	TTTCACGTCA	TACAGTCGTT	ATCCCCGTAT	1050
	GATGGCCGCT	GCAGGTATTG	CACTTCCCGA	ACTGATTGAC	CGCTTGATCG	1100
	TATTAGCGTT	AAAGGGGTGA	TAAGCATGGA	AATAGGATTT	ACTTTTTTTAG	1150
	ATGAAATAGT	ACACGGTGTT	CGTTGGGACG	CTAAATATGC	CACTTGGGAT	1200
	AATTTACCG	GAAAACCGGT	TGACGGGTAT	GAAAGTAAAT	CGCATTGTAG	1250
45	GGACATTTCGA	ATT				1263

## 2) INFORMATION FOR SEQ ID NO: 1054

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1232 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*

(B) STRAIN: R610

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1054

```

5  CGGCCTATTA TNCCGAGCAA GCGTTGCGTG ATACCGTTGA AAAAACCATT 50
   AAAAAGTGGT TGGATTTTGA AAGGAGACAG GAGCATGAAT AGAATAAAAG 100
   TTGCAATACT GTTTGGGGGT TGCTCAGAGG AGCATGACGT ATCGGTAAAA 150
   TCTGCAATAG AGATAGCCGC TAACATTAAT AAAGAAAAAT ACGAGCCGTT 200
   ATACATTGGA ATTACGAAAT CTGGTGTATG GAAAATGTGC GAAAAACCTT 250
10  GCGCGGAATG GGAACGAC AATTGCTATT CAGCTGTACT CTCGCCGGAT 300
   AAAAAAATGC ACGGATTACT TGTAAAAAG AACCATGAAT ATGAAATCAA 350
   CCATGTTGAT GTAGCATTTT CAGCTTTGCA TGGCAAGTCA GGTGAAGATG 400
   GATCCATACA AGGTCTGTTT GAATTGTCCG GTATCCCTTT TGTAGGCTGC 450
   GATATTCAAA GCTCAGCAAT TTGTATGGAC AAATCGTTGA CATAATCGT 500
15  TGCAGAAAAAT GCTGGGATAG CTACTCCCGC CTTTGGGGT ATTAATAAAG 550
   ATGATAGGCC GGTGGCAGCT ACGTTTACCT ATCCTGTTTT TGTAAAGCCG 600
   GCGCGTTCAG GCTCATCCTT CCGTGTGAAA AAAGTCAATA GCGCGGACGA 650
   ATTGGACTAC GCAATTGAAT CGGCAAGACA ATATGACAGC AAAATCTTAA 700
   TTGAGCAGGC TGTTTCGGGC TGTGAGGTCG GTTGTGCGGT ATTGGGAAAC 750
20  AGTGCCGCGT TAGTTGTTGG CGAGGTGGAC CAAATCAGGC TGCAGTACGG 800
   AATCTTTCGT ATTCATCAGG AAGTCGAGCC GGAAAAAGGC TCTGAAAACG 850
   CAGTTATAAC CGTTCCCGCA GACCTTTCAG CAGAGGAGCG AGGACGGATA 900
   CAGGAAACGG CAAAAAAAT ATATAAAGCG CTCGGCTGTA GAGGTCTAGC 950
   CCGTGTGGAT ATGTTTTTAC AAGATAACGG CCGCATTGTA CTGAACGAAG 1000
25  TCAATACTCT GCCCGGTTTC ACGTCATACA GTCGTTATCC CCGTATGATG 1050
   GCCGCTGCAG GTATTGCACT TCCCGAACTG ATTGACCGCT TGATCGTATT 1100
   AGCGTTAAAG GGGTGATAAG CATGGAAATA GGATTTACTT TTTTAGATGA 1150
   AATAGTACAC GGTGTTCTGT GGGACGCTAA ATATGCCACT TGGGATAATT 1200
   TCACCGGAAA ACCGGTTGAC GGTATAAGT AA 1232
30

```

2) INFORMATION FOR SEQ ID NO: 1055

```

35  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 1218 bases
      (B) TYPE: Nucleic acid
      (C) STRANDEDNESS: Double
      (D) TOPOLOGY: Linear
40  (ii) MOLECULE TYPE: Genomic DNA
      (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Enterococcus gallinarum
          (B) STRAIN: R684
45

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1055

```

50  TACCGAGCAA GCGTTGCGTG ATACCGTTGA AAAAACCATT AAAAAGTGGT 50
   TGGATTTTGA AAGGAGACAG GAGCATGAAT AGAATAAAAG TTGCAATACT 100
   GTTTGGGGGT TGCTCAGAGG AGCATGACGT ATCGGTAAAA TCTGCAATAG 150
   AGATAGCCGC TAACATTAAT AAAGAAAAAT ACGAGCCGTT ATACATTGGA 200
   ATTACGAAAT CTGGTGTATG GAAAATGTGC GAAAAACCTT GCGCGGAATG 250
   GGAAAAACGAC AATTGCTATT CAGCTGTACT CTCGCCGGAT AAAAAATGC 300
55  ACGGATTACT TGTAAAAAG AACCATGAAT ATGAAATCAA CCATGTTGAT 350
   GTAGCATTTT CAGCTTTGCA TGGCAAGTCA GGTGAAGATG GATCCATACA 400
   AGGTCTGTTT GAATTGTCCG GTATCCCTTT TGTAGGCTGC GATATTCAAA 450
   GCTCAGCAAT TTGTATGGAC AAATCGTTGA CATAATCGT TGCAGAAAAAT 500
   GCTGGGATAG CTACTCCCGC CTTTGGGGT ATTAATAAAG ATGATAGGCC 550
60  GGTGGCAGCT ACGTTTACCT ATCCTGTTTT TGTAAAGCCG GCGCGTTCAG 600

```

	GCTCATCCTT	CGGTGTGAAA	AAAGTCAATA	GCGCGGACGA	ATTGGACTAC	650
	GCAATTGAAT	CGGCAAGACA	ATATGACAGC	AAAATCTTAA	TTGAGCAGGC	700
	TGTTTCGGGC	TGTGAGGTCG	GTTGTGCGGT	ATTGGGAAAC	AGTGCCGCGT	750
	TAGCTGTTGG	CGAGGTGGAC	CAAATCAGGC	TGCAGTACGG	AATCTTTTCGT	800
5	ATTCATCAGG	AAGTCGAGCC	GGAAAAAGGC	TCTGAAAACG	CAGTTATAAC	850
	CGTTCCCGCA	GACCTTTTCAG	CAGAGGAGCG	AGGACGGATA	CAGGAAACGG	900
	CAAAAAAAT	ATATAAAGCG	CTCGGCTGTA	GAGGTCTAGC	CCGTGTGGAT	950
	ATGTTTTTAC	AAGATAACGG	CCGCATTGTA	CTGAACGAAG	TCAATACTCT	1000
	GCCCGGTTTC	ACGTCATACA	GTCGTTATCC	CCGTATGATG	GCCGCTGCAG	1050
10	GTATTGCACT	TCCCGAAGTG	ATTGACCGCT	TGATCGTATT	AGCGTTAAAG	1100
	GGGTGATAAG	CATGGAAATA	GGATTTACTT	TTTTAGATGA	AATAGTACAC	1150
	GGTGTTCGTT	GGGACGCTAA	ATATGCCACT	TGGGATAATT	TCACCGGAAA	1200
	ACCGGTTGAC	GGTTAGAA				1218

15

## 2) INFORMATION FOR SEQ ID NO: 1056

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1265 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Enterococcus faecium*
- (B) STRAIN: R688
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1056

	AATCACACCG	CATACGGCCT	ATTATACCGA	GCAAGCGTTG	CGTGATACCG	50
	TTGAAAAAAC	CATTAAAAAC	TGTTTGATG	TTGAAAGGAG	ACAGGAGCAT	100
35	GAATAGAATA	AAAGTTGCAA	TACTGTTTGG	GGGTTGCTCA	GAGGAGCATG	150
	ACGTATCGGT	AAAATCTGCA	ATAGAGATAG	CCGCTAACAT	TAATAAAGAA	200
	AAATACGAGC	CGTTATACAT	TGGAATTACG	AAATCTGGTG	TATGAAAAAT	250
	GTGCGAAAAA	CCTTGCGCGG	AATGGGAAAA	CGACAATTGC	TATTCAGCTG	300
	TACTCTCGCC	GGATAAAAAA	ATGCACGGAT	TACTTGTTAA	AAAGAACCAT	350
40	GAATATGAAA	TCAACCATGT	TGATGTAGCA	TTTTTCAGCTT	TGCATGGCAA	400
	GTCAGGTGAA	GATGGATCCA	TACAAGGTCT	GTTTGAATTG	TCCGGTATCC	450
	CTTTTGTAGG	CTGCGATATT	CAAAGCTCAG	CAATTTGTAT	GGACAAATCG	500
	TTGACATACA	TCGTTGCGAA	AAATGCTGGG	ATAGCTACTC	CCGCCTTTTG	550
	GGTTATTAAT	AAAGATGATA	GGCCGGTGCG	AGCTACGTTT	ACCTATCCTG	600
45	TTTTTGTTAA	GCCGGCGCGT	TCAGGCTCAT	CCTTCGGTGT	GAAAAAAGTC	650
	AATAGCGCGG	ACGAATTGGA	CTACGCAATT	GAATCGGCAA	GACAATATGA	700
	CAGCAAAATC	TTAATTGAGC	AGGCTGTTTC	GGGCTGTGAG	GTCGGTTGTG	750
	CGGTATTGGG	AAACAGTGCC	GCGTTAGCTG	TTGGCGAGGT	GGACCAAATC	800
	AGGCTGCAGT	ACGGAATCTT	TCGTATTCAT	CAGGAAGTCG	AGCCGGAAAA	850
50	AGGCTCTGAA	AACGCAGTTA	TAACCGTTCC	CGCAGACCTT	TCAGCAGAGG	900
	AGCGAGGACG	GATACAGGAA	ACGGCAAAAA	AAATATATAA	AGCGCTCGGC	950
	TGTAGAGGTC	TAGCCCGTGT	GGATATGTTT	TTACAAGATA	ACGGCCGCAT	1000
	TGTACTGAAC	GAAGTCAATA	CTCTGCCCGG	TTTCACGTCA	TACAGTCGTT	1050
	ATCCCCGTAT	GATGGCCGCT	GCAGGTATTG	CACTTCCCGA	ACTGATTGAC	1100
55	CGCTTGATCG	TATTAGCGTT	AAAGGGGTGA	TAAGCATGGA	AATAGGATTT	1150
	ACTTTTTTAG	ATGAAATAGT	ACACGGTGTT	CGTTGGGACG	CTAAATATGC	1200
	CACTTGGGAT	AATTTACCCG	GAAAACCGGT	TGACGGTTAT	GAAGTAAATC	1250
	GCATTGTAGG	GACAT				1265

60



## 2) INFORMATION FOR SEQ ID NO: 1057

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1269 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus flavescens*  
 (B) STRAIN: R689

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1057

TCACACCGCA	TACGGCCTAT	TATACCGAGC	AAGCGTTGCG	TGATACCGTT	50
GAAAAAACCA	TTAAAAACTG	TTTGGATTTT	GAAAGGAGAC	AGGAGCATGA	100
ATAGAATAAA	AGTTGCAATA	CTGTTTGGGG	GTTGCTCAGA	GGAGCATGAC	150
GTATCGGTAA	AATCTGCAAT	AGAGATAGCC	GCTAACATTA	ATAAAGAAAA	200
ATACGAGCCG	TTATACATTG	GAATTACGAA	ATCTGGTGTA	TGGAAAATGT	250
GCGAAAAACC	TTGCGCGGAA	TGGGAAAACG	ACAATTGCTA	TTCAGCTGTA	300
CTCTCGCCGG	ATAAAAAAAT	GCACGGATTA	CTTGTTAAAA	AGAACCATGA	350
ATATGAAATC	AACCATGTTG	ATGTAGCATT	TTCAGCTTTG	CATGGCAAGT	400
CAGGTGAAGA	TGGATCCATA	CAAGGTCTGT	TTGAATTGTC	CGGTATCCCT	450
TTTGTAGGCT	GCGATATTCA	AAGCTCAGCA	ATTTGTATGG	ACAAATCGTT	500
GACATACATC	GTTGCGAAAA	ATGCTGGGAT	AGCTACTCCC	GCCTTTTGGG	550
TTATTAATAA	AGATGATAGG	CCGGTGGCAG	CTACGTTTAC	CTATCCTGTT	600
TTTGTTAAGC	CGGCGCGTTC	AGGCTCATCC	TTCGGTGTGA	AAAAAGTCAA	650
TAGCGCGGAC	GAATTGGACT	ACGCAATTGA	ATCGGCAAGA	CAATATGACA	700
GCAAAATCTT	AATTGAGCAG	GCTGTTTCGG	GCTGTGAGGT	CGGTTGTGCG	750
GTATTGGGAA	ACAGTGCCGC	GTTAGCTGTT	GGCGAGGTGG	ACCAAATCAG	800
GCTGCAGTAC	GGAATCTTTC	GTATTCATCA	GGAAGTCGAG	CCGAAAAAAG	850
GCTCTGAAAA	CGCAGTTATA	ACCGTTCCCG	CAGACCTTTC	AGCAGAGGAG	900
CGAGGACGGA	TACAGGAAAC	GGCAAAAAAA	ATATATAAAG	CGCTCGGCTG	950
TAGAGGTCTA	GCCCGTGTGG	ATATGTTTTT	ACAAGATAAC	GGCCGCATTG	1000
TACTGAACGA	AGTCAATACT	CTGCCCCGTT	TCACGTCATA	CAGTCGTTAT	1050
CCCCGTATGA	TGGCCGCTGC	AGGTATTGCA	CTTCCCGAAC	TGATTGACCG	1100
CTTGATCGTA	TTAGCGTTAA	AGGGGTGATA	AGCATGGAAA	TAGGATTTAC	1150
TTTTTTTAGAT	GAAATAGTAC	ACGGTGTTTCG	TTGGGACGCT	AAATATGCCA	1200
CTTGGGATAA	TTTCACCGGA	AAACCGGTTG	ACGGTTATGA	AGTAAATCGC	1250
ATTGTAGGGA	CATACGAAT				1269

## 2) INFORMATION FOR SEQ ID NO: 1058

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1169 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus gallinarum*  
 (B) STRAIN: R420

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1058

	CAAATTTTCT	TTTCTTTTCC	TAGGTACACT	GAATGTAACC	TTAAAAGAAA	50
	AAAGGAAAGG	AAGAAAATGA	TGAAAAAAT	TGCCGTTTTA	TTTGGAGGGA	100
5	ATTCTCCAGA	ATACTCAGTG	TCACTAACCT	CAGCAGCAAG	TGTGATCCAA	150
	GCTATTAACC	CGCTGAAATA	TGAAGTAATG	ACCATTGGCA	TCGCACCAAC	200
	AATGGATTGG	TATTGGTATC	AAGGAAACCT	CGCGAATGTT	CGCAATGATA	250
	CTTGGCTAGA	AGATCACAAA	AACTGTCACC	AGCTGACTTT	TTCTAGCCAA	300
	GGATTTATAT	TAGGAGAAAA	ACGAATCGTC	CCTGATGTCC	TCTTTCCAGT	350
10	CTTGCATGGG	AAGTATGGCG	AGGATGGCTG	TATCCAAGGA	CTGCTTGAAC	400
	TAATGAACCT	GCCTTATGTT	GGTTGCCATG	TCGCTGCCTC	CGCATTATGT	450
	ATGAACAAAT	GGCTCTTGCA	TCAACTTGCT	GATACCATGG	GAATCGCTAG	500
	TGCTCCCACT	TTGCTTTTAT	CCCGCTATGA	AAACGATCCT	GCCACAATCG	550
	ATCGTTTTAT	TCAAGACCAT	GGATTCCCGA	TCTTTATCAA	GCCGAATGAA	600
15	GCCGGTTCTT	CAAAAGGGAT	CACAAAAGTA	ACTGACAAAA	CAGCGCTCCA	650
	ATCTGCATTA	ACGACTGCTT	TTGCTTACGG	TTCTACTGTG	TTGATCCAAA	700
	AGGCGATAGC	GGGTATTGAA	ATTGGCTGCG	GCATCTTAGG	AAATGAGCAA	750
	TTGACGATTG	GTGCTTGTGA	TGCGATTTCT	CTTGTCGACG	GTTTTTTTGA	800
	TTTTGAAGAG	AAATACCAAT	TAATCAGCGC	CACGATCACT	GTCCCAGCAC	850
20	CATTGCCTCT	CGCGCTTGAA	TCACAGATCA	AGGAGCAGGC	ACAGCTGCTT	900
	TATCGAAACT	TGGGATTGAC	GGGTCTGGCT	CGAATCGATT	TTTTTCGTAC	950
	CAATCAAGGA	GCGATTTATT	TAAACGAAAT	CAACACCATG	CCGGGATTTA	1000
	CTGGGCACTC	CCGCTACCCA	GCTATGATGG	CGGAAGTCGG	GTTATCCTAC	1050
	GAAATATTAG	TAGAGCAATT	GATTGCACTG	GCAGAGGAGG	ACAAACGATG	1100
25	AACACATTAC	AATTGATCAA	TAAAAACCAT	CCATTGAAAA	AAAATCAAGA	1150
	GCCCCCGCAC	TTAGTGCTA				1169

## 30 2) INFORMATION FOR SEQ ID NO: 1059

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1166 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 40 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus gallinarum*  
 (B) STRAIN: R631

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1059

45	CAAATTTTCT	TTTCTTTTCC	TAGGTACACT	GAATGTAACC	TTAAAAGAAA	50
	AAAGGAAAGG	AAGAAAATGA	TGAAAAAAT	TGCCGTTTTA	TTTGGAGGGA	100
	ATTCTCCAGA	ATACTCAGTG	TCACTAGCCT	CAGCAGCAAG	TGTGATCCAA	150
	GCTATTGACC	CGCTGAAATA	TGAAGTAATG	ACCATTGGCA	TCGCACCAAC	200
50	AATGGATTGG	TATTGGTATC	AAGGAAACCT	CGCGAATGTT	CGCAATGATA	250
	CTTGGCTAGA	AGATCACAAA	AACTGTCACC	AGCTGACTTT	TTCTAGCCAA	300
	GGATTTATAT	TAGGAGAAAA	ACGAATCGTC	CCTGATGTCC	TCTTTCCAGT	350
	CTTGCATGGG	AAGTATGGCG	AGGATGGCTG	TATCCAAGGA	CTGCTTGAAC	400
	TAATGAACCT	GCCTTATGTT	GGTTGCCATG	TCGCTGCCTC	CGCATTATGT	450
55	ATGAACAAAT	GGCTCTTGCA	TCAACTTGCT	GATACCATGG	GAATCGCTAG	500
	TGCTCCCACT	TTGCTTTTAT	CCCGCTATGA	AAACGATCCT	GCCACAATCG	550
	ATCGTTTTAT	TCAAGACCAT	GGATTCCCGA	TCTTTATCAA	GCCGAATGAA	600
	GCCGGTTCTT	CAAAAGGGAT	CACAAAAGTA	ACTGACAAAA	CAGCGCTCCA	650
	ATCTGCATTA	ACGACTGCTT	TTGCTTACGG	TTCTACTGTG	TTGATCCAAA	700
60	AGGCGATAGC	GGGTATTGAA	ATTGGCTGCG	GCATCTTAGG	AAATGAGCAA	750

10  
 5  
 TTGACGATTG GTGCTTGTGA TGC GATTCTCT CTTGTCGACG GTTTTTTTTGA 800  
 TTTTGAAGAG AAATACCAAT TAATCAGCGC CACGATCACT GTCCCAGCAC 850  
 CATTGCCTCT CGCGCTTGAA TCACAGATCA AGGAGCAGGC ACAGCTGCTT 900  
 TATCGAAACT TGGGATTGAC GGGTCTGGCT CGAATCGATT TTTTCGTCAC 950  
 CAATCAAGGA GCGATTTATT TAAACGAAAT CAACACCATG CCGGGATTTA 1000  
 CTGGGCACTC CCGCTACCCA GCTATGATGG CGGAAGTCGG GTTATCCTAC 1050  
 GAAATATTAG TAGAGCAATT GATTGCACTG GCAGAGGAGG ACAAACGATG 1100  
 AACACATTAC AATTGATCAA TAAAAACCAT CCATTGAAAA AAAATCAAGA 1150  
 GCCCCCGCAC TTAGTG 1166

## 2) INFORMATION FOR SEQ ID NO: 1060

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1028 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear  
 20 (ii) MOLECULE TYPE: Genomic DNA  
 (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Enterococcus casseliflavus*  
 25 (B) STRAIN: ATCC 25788  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1060

30 AACATGAAAA AAATCGCCTT ATTTTGGAG GCAATTCACC GGAATACACC 50  
 GTTTCCTTTAG CTTTCAGCAAC TAGCGCAATC GAAGCACTCC AATCATCTCC 100  
 CTATGACTAC GACCTCTCTT TGATCGGGAT CGCCCCAGAT GCTATGGATT 150  
 GGTACTTGTA TACAGGAGAA CTGGAAAACA TCCGACAAGA CACGTGGTTG 200  
 TTGGATACGA AACATAAACA GAAAATACAG CCGCTATTCG AAGGAAACGG 250  
 CTTTGGCTA AGTGAAGAGC AGCAAACGTT GGTACCTGAT GTTTTATTTC 300  
 35 CCATTATGCA TGGCAAATAC GGGGAAGATG GCAGTATCCA AGGATTGTTT 350  
 GAATTGATGA AGCTGCCTTA TGTAGGCTGC GGGGTGGCAG GTTCTGCCTT 400  
 ATGTATGAAC AAATGGCTGC TGCATCAAGC TGCAGCAGCC ATTGGCGTAC 450  
 AAAGTGCTCC TACGATTCTC TTGACAAATC AAGCCAACCA GCAAGAACAA 500  
 ATCGAAGCTT TTATCCAGAC CCATGGCTTC CCAGTTTCT TTAAGCCTAA 550  
 40 TGAAGCGGGC TCCTCAAAAG GGATCACTAA AGTCACCTGC GTTGAAGAAA 600  
 TCGCTTCTGC CTTAAAAGAA GCCTTTACTT ATTGTTCCGC AGTGCTCCTA 650  
 CAAAAAATA TTGCCGGTGT TGAGATCGGT TGCGGTATTT TGGGCAACGA 700  
 CTCTTTGACT GTCGGTGCTT GTGACGCCAT TTCATTAGTA GACGGCTTTT 750  
 TCGATTTTGA AGAAAAGTAC CAGCTGATCA GCGCCAAAAT CACCGTCCCT 800  
 45 GCGCCATTGC CTGAAACGAT TGAAACCAAG GTCAAAGAAC AAGCTCAGCT 850  
 GCTCTATCGT AGTCTTGGTC TTAAAGGTCT TGCTCGCATC GACTTTTTTTG 900  
 TCACGGAGCG AGGAGAACTA TACTTGAATG AAATCAATAC TATGCCGGGC 950  
 TTTACGAGTC ACTCCCGCTA TCCTGCCATG ATGGCAGCGG TCGGCTTATC 1000  
 50 CTATCAAGAA CTACTACAAA AACTGCTT 1028

## 2) INFORMATION FOR SEQ ID NO: 1061

55 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1030 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 60 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus casseliflavus*  
(B) STRAIN: R689

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1061

```

10  AATATGAAAA AAATCGCCTT ATTTTGGAG GCAATTCACC GGAGTACGCC      50
    GTTTCCTTTAG CCTCAGCAAC TAGCGCAATC GAAGCACTCC AATCATCTCC      100
    CGATGACTAT  GACCTCTCTT TGATCGGGAT CGCCCCAGAT GCTATGGATT      150
    GGTATTTGTA  TACAGGAGAA CTGGAAAACA TCCGACAAGA CACGTGGTTG      200
    TTGGATACGA  AACATAAACA GAAAATCCAG CCGCTTTTGT AAGGAAACGG      250
    CTTTTGGCTA  AGTGAAGAGC AACAAACGTT GGTTCCTGAT GTTTTATTTT      300
15  CCATTATGCA  TGGCAAATAC GGGGAAGATG GCAGTATCCA AGGATTGTTT      350
    GAATTGATGA  AACTACCTTA TGTAGGCTGC GGGGTGGCAG CCTCTGCCTT      400
    ATGTATGAAC  AAATGGCTGC TGCATCAAGC AGCAGAAGCG ATTGGCGTAC      450
    AAAGTGCTCC  TACGATTCTC TTGACAAATC AAGCCAACCA GCAAGATCAA      500
    ATCGAAGCTT  TTATCCAGAC CCATGGCTTT CCGGTTTTTT TTAAGCCTAA      550
20  TGAAGCGGGC  TCCTCAAAAG GGATCACTAA AGTCACCTGC GTTGAAGAAA      600
    TCGCTTCTGC  CTTAAAAGAA GCCTTTACTT ATTGTTTCAGC AGTGCTCCTA      650
    CAAAAAATA  TTGCCGGTGT TGAGATCGGT TGCGGTATTT TGGGCAACGA      700
    CTCTTTGACT  GTCGGTGCTT GTGACGCTAT TTCATTAGTA GACGGCTTTT      750
    TCGATTTTGA  AGAAAAGTAC CAGCTGATCA GCGCCAAGAT CACCGTTCCT      800
25  GCACCATTGC  CTGAAACGAT TGAAACCAAG GTCAAAGAAC AAGCTCAGCT      850
    GCTCTATCGT  AGTCTTTGTC TTAAAGGTCT TGCTCGCATC GACTTTTTTT      900
    TCACGGATCA  AGGAGAACTA TACTTGAATG AAATCAATAC TATGCCGGGC      950
    TTTACGAGTC  ACTCCCGCTA TCCTGCCATG ATGGCAGCGA TCGGCTTATC     1000
    CTATCAAGAA  CTACTACAAA AACTGCTTGT      1030
30

```

2) INFORMATION FOR SEQ ID NO: 1062

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1031 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus casseliflavus*  
(B) STRAIN: R754

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1062

```

50  AAACATGAAA AAAATCGCCA TTATTTTGG AGGCAATTCA CCGGAATACA      50
    CCGTTTCTTT AGCTTCAGCA ACTAGCGCAA TCGAAGCACT CCAATCATCT      100
    CCCTATGACT  ACGACCTCTC TTTGATCGGG ATCGCCCCAG ATGCTATGGA      150
    TTGGTACTTG  TATACAGGAG AACTGGAAAA CATCCGACAA GACACGTGGT      200
    TGTTGGATAC  GAAACATAAA CAGAAAATAC AGCCGCTATT CGAAGGAAAC      250
    GGCTTTTGGC  TAAGTGAAGA GCAGCAAACG TTGGTACCTG ATGTTTTATT      300
55  TCCCATTATG  CATGGCAAAAT ACGGGGAAGA TGGCAGTATC CAAGGATTGT      350
    TTGAATTGAT  GAAGCTGCCT TATGTAGGCT GCGGGGTGGC AGGTTCTGCC      400
    TTATGTATGA  ACAAATGGCT GCTGCATCAA GCTGCAGCAG CCATTGGCGT      450
    ACAAAGTGCT  CCTACGATTC TCTTGACAAA TCAAGCCAAC CAGCAAGAAC      500
    AAATCGAAGC  TTTTATCCAG ACCCATGGCT TCCCAGTTTT CTTTAAGCCT      550
60  AATGAAGCGG  GCTCCTCAAA AGGGATCACT AAAGTCACCT GCGTTGAAGA      600

```

```

AATCGCTTCT GCCTTAAAAAG AAGCCTTTAC TTATTGTTCC GCAGTGCTCC 650
TACAAAAAAA TATTGCCGGT GTTGAGATCG GTTGCGGTAT TTTGGGCAAC 700
GACTCTTTGA CTGTCGGTGC TTGTGACGCC ATFTCATTAG TAGACGGCTT 750
TTTCGATTTT GAAGAAAAGT ACCAGCTGAT CAGCGCCAAA ATCACCGTCC 800
5 CTGCGCCATT GCCTGAAACG ATTGAAACCA AGGTCAAAGA ACAAGCTCAG 850
CTGCTCTATC GTAGTCTTGG TCTTAAAGGT CTTGCTCGCA TCGACTTTTT 900
TGTCACGGAG CGAGGAGAAC TATACTTGAA TGAAATCAAT ACTATGCCGG 950
GCTTTACGAG TCACTCCCGC TATCCTGCCA TGATGGCAGC GGTCGGCTTA 1000
10 TCCTATCAAG AACTACTACA AAAACTGCTT G 1031

```

## 2) INFORMATION FOR SEQ ID NO: 1063

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1030 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
20 (ii) MOLECULE TYPE: Genomic DNA
    (vi) ORIGINAL SOURCE:
    (A) ORGANISM: Enterococcus casseliflavus
25 (B) STRAIN: R775

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1063

```

AACATGAAAA AAATCGCCAT TATTTTTGGA GGCAATTCAC CGGAATACAC 50
30 CGTTTCTTTA GCTTCAGCAA CTAGCGCAAT CGAAGCACTC CAATCATCTC 100
CCTATGACTA CGACCTCTCT TTGATCGGGA TCGCCCCAGA TGCTATGGAT 150
TGGTACTTGT ATACAGGAGA ACTGGAAAAC ATCCGACAAG ACACGTGGTT 200
GTTGGATACG AAACATAAAC AGAAAATACA GCCGCTATTT GAAGGAAACG 250
GCTTTTGGCT AAGTGAAGAG CAGCAAACGT TAGTACCTGA TATTTTATTT 300
35 CCCATTATGC ATGGCAAATA CGGGGAAGAT GGCAGTATCC AAGGATTGTT 350
TGAATTGATG AAATACCTT ATGTAGGTTG CGGGGTGGCA GGTTCGCTT 400
TATGTATGAA CAAATGGCTG CTGCATCAAG CTGCAGCAGC CATTGGCGTA 450
CAAAGTGCTC CTACGATTCT CTTGACAAAT CAAGCCAACC AGCAAGAACA 500
AATCGAAGCT TTTATCCAGA CCCATGGCTT CCCAGTTTTC TTTAAGCCTA 550
40 ATGAAGCGGG CTCTTCAAAA GGGATCACTA AAGTCACCTG CGTTGAAGAA 600
ATCGCTTCTG CCTTAAAAAA AGCCTTTACT TATTGTTCCG CAGTGCTCCT 650
ACAAAAAAT ATTGCCGGTG TTGAGATCGG TTGCGGTATT TTTGGGCAACG 700
ACTCTTTGAC TGTCGGTGCT TGTGACGCCA TTTCATTAGT AGACGGCTTT 750
TTCGATTTTG AAGAAAAGTA CCAGCTGATC AGCGCCAAAA TCACCGTCCC 800
45 TGCGCCATTG CCTGAAACGA TTGAAACCAA GGTCAAAGAA CAAGCTCAGC 850
TGCTCTATCG TAGTCTTGGT CTTAAAGGTC TTGCTCGCAT CGACTTTTTT 900
GTCACGGATC AAGGAGAACT AACTTTGAAT GAAATCAATA CTATGCCGGG 950
CTTTACGAGT CACTCCCGTT ATCCTGCCAT GATGGCAGCG GTCGGCTTAT 1000
50 CCTATCAAGA ACTACTACAA AAATGCTTG 1030

```

## 2) INFORMATION FOR SEQ ID NO: 1064

```

55 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1032 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
60

```

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus flavescens*

(B) STRAIN: ATCC 49996

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1064

```

10 AAACATGAAA AAAATCGCCA TTATTTTGG AGGCAATTCA CCGGAATACA      50
   CCGTTTCTTT AGCTTCAGCA ACTAGCGCAA TCGAAGCACT CCAATCATCT      100
   CCCTATGACT ACGACCTCTC TTTGATCGGG ATCGCCCCAG ATGCTATGGA      150
   TTGGTACTTG TATACAGGAG AACTGGAAAA CATCCGACAA GACACGTGGT      200
   TGTGATGATC GAAACAGAAA CAGAAAATAC AGCCGCTATT CGAAGGAAAC      250
   GGCTTTTGGT TAAGTGAAGA GCAGCAAACG TTGGTACCTG ATGTTTATT      300
15 TCCCATTATG CATGGCAAAT ACGGGGAAGA TGGCAGTATC CAAGGATTGT      350
   TTGAATTGAT GAAGCTACCT TATGTAGGCT GCGGGGTGGC AGGTTCTGCC      400
   TTATGTATGA ACAAATGGTT GCTGCATCAA GCTGCAGCAG CCATTGGCGT      450
   ACAAAGTGCT CCTACGATTC TCTTGACAAA TCACGCCAAC CAGCAAGAAC      500
   AAATCGAAGC TTTTATCCAG ACCCATGGCT TTCCAGTTT CTTTAAGCCT      550
20 AATGAAGCGG GTTCTCTCAA AGGGATCACT AAAGTCACCT GCGTTGAAGA      600
   AATCGCTTCT GCCTTAAAG AAGCCTTTAC TTATTGTTCC GCAGTGCTCC      650
   TACAAAAAAT CATTGCCGGT GTTGAGATCG GTTGCGGTAT TTTGGGCAAC      700
   GACTCTTTGA CTGTGCGTGC TTGTGACGCC ATTCATTAG TAGACGGCTT      750
   TTTTCGATTT GAAGAAAAGT ACCAGCTGAT CAGCGCCAAA ATCACCGTCC      800
25 CTGCGCCATT GCCTGAAACG ATTGAACTA AGGTCAAAGA ACAAGCTCAG      850
   CTGCTCTATC GTAGTCTTGG ACTTAAAGGT CTTGCTCGCA TCGACTTTTT      900
   TGTCACGGAT CAAGGAGAAC TATACTTGAA TGAAATCAAT ACTATGCCGG      950
   GCTTTACGAG TCACTCCCGC TATCTGCCA TGATGGCAGC GGTGCGGTTA     1000
30 TCCTATCAAG AACTACTACA AAAACTACTT GT                               1032

```

## 2) INFORMATION FOR SEQ ID NO: 1065

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1034 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus flavescens*

(B) STRAIN: R758

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1065

```

50 AAAAACATGA AAAAAATCGC CATTATTTTT GGAGGCAATT CACCGGAATA      50
   CACCGTTTCT TTAGCTTCAG CAACTAGCGC AATCGAAGCA CTCCAATCAT      100
   CTCCCTATGA CTACGACCTC TCTTTGATCG GGATCGCCCC AGATGCTATG      150
   GATTGGTACT TGTATACAGG AGAACTGGAA AACATCCGAC AAGACACGTG      200
   GTTGTGATG ACGAAACATA AACAGAAAAT ACAGCCGCTA TTCGAAGGAA      250
   ACGGCTTTTG GCTAAGTGAA GAGCAGCAAA CGTTGGTACC TGATGTTTTA      300
55 TTTCCCATTA TGCATGGCAA ATACGGGGAA GATGGCAGTA TCCAAGGATT      350
   GTTTGAATTG ATGAAGCTGC CTTATGTAGG CTGCGGGGTG GCAAGTTCTG      400
   CCTATGTAT GAACAAATGG CTGCTGCATC AAGCTGCAGC AGCCATTGGC      450
   GTACAAAGTG CTCCTACGAT TCTCTTGACA AATCAAGCCA ACCAGCAAGA      500
   ACAAATCGAA GCTTTTATCC AGACCCATGG CTTTCCAGTT TTCTTTAAGC      550
60 CTAATGAAGC GGGCTCCTCA AAAGGGATCA CTAAAGTCAC CTGCGTTGAA      600

```

```

GAAATCGCTT CTGCCTTAAA AGAAGCCTTT ACTTATTGTT CCGCAGTGCT 650
CCTACAAAAA AATATTGCCG GTGTTGAGAT CGGTTGCGGT ATTTTGGGCA 700
ACGACTCTTT GACTGTCGGT GCTTGTGACG CCATTTTCATT AGTAGACGGC 750
TTTTTCGATT TTGAAGAAAA GTACCAGCTG ATCAGCGCCA AAATCACCGT 800
5 CCCTGCGCCA TTGCCTGAAA CGATTGAAAC CAAGGTCAAA GAACAAGCTC 850
AGCTGCTCTA TCGTAGTCTT GGTCTTAAAG GTCTTGCTCG CATCGACTTT 900
TTTGTACACG ATCAAGGAGA ACTATACTTG AATGAAATCA ATACTATGCC 950
GGGCTTTACG AGTCACTCCC GCTATCCTGC CATGATGGCA GCGGTCGGCT 1000
TATCCTATCA AGAACTACTA CAAAACTGC TTGT 1034
10

```

## 2) INFORMATION FOR SEQ ID NO: 1066

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 1012 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
20
(ii) MOLECULE TYPE: Genomic DNA
(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Enterococcus flavescens
25    (B) STRAIN: R760
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1066

```

```

CATGAAAAAA ATCGCCATTA TTTTGGAGG CAATTCACCG GAATACACCG 50
30 TTTCTTTAGC CTCAGCAACT AGCGCAATCG AAGCACTCCA ATCATCTCCC 100
TATGATTACG ACCTCTCTTT GATCGGGATC GCCCCAGATG CTATGGATTG 150
GTACTTGAT ATCAGGAGAAC TGGAAAACAT CCGACAAGAC ACGTGGTTGT 200
TGGATACGAA ACATACACAG AAAATCCAGC CACTTTTTGA AGGAAACGGC 250
TTTTGGATAA GTGAAGCGCA GCAAACGTTG GTACCTGATG TTTTATTTCC 300
35 CATTATGCAT GGTAAATACG GGAAGATGG CAGTATCCAA GGATTGTTTG 350
AATTGATGAA GCTGCCTTAT GTAGGCTGTG GGGTGGCAGC CTCTGCCTTA 400
TGTATGAACA AATGGTTATT GCATCAAGCA GCAGCAGCGA TTGGCGTACA 450
AAGCGCTCCT ACGATTCTCT TGACAAATCA AGCCAACCAG CAAAGACAAA 500
TCGAAGCCTT TATCCAGACC ATGGGCTTTC CAGTTTTCTT TAAGCCTAAT 550
40 GAAGCGGGCT CCTCAAAAGG GATCACAAA GTAACCTGTG TTGAAGAAAT 600
CGCTCCTGCC TTGAAGGAAG CCTTCGCTTA TTGTTCCGCA GTGCTCTTAC 650
AAAAAATAT CGCTGGCGTT GAGATTGGTT GCGGTATCTT AGGCAACGAC 700
TCTTTGACTG TCGGTGCTTG TGACGCTATT TCATTAGTAG ACGGCTTTTT 750
CGATTTTGAA GAAAAGTACC AGTTGATCAG CGCCAAAATC ACCGTTCTTG 800
45 CGCCATTGCC TGAAACGATT GAAACCAAAG TCAAAGAACA AGCTCAGCTG 850
CTCTATCACA GTCTTGGTCT TAAAGGACTT GCTCGCATCG ACTTTTTTGT 900
CACGGATCAA GGAGAACTAT ACTTGAATGA AATCAATACT ATGCCGGGCT 950
TTACGAGTCA CTCCCGCTAT CCTGCCATGA TGGCAGCGGT CGGCTTATCC 1000
TATCAAGAAT TA 1012
50

```

## 2) INFORMATION FOR SEQ ID NO: 1067

```

55 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 721 bases
    (B) TYPE: Nucleic acid
    (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
60

```

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*

(B) STRAIN: R481

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1067

```

10 CTTACGCTTT ATCGATTAGA CACGGGNAGC TTGTCCAATG GGRAGCCGAT      50
   TTGATTTTAT GGATGAACGC TCTCATCATG CGGCAAATGG AATATCATGC      100
   AATGAAGCGC AAAATCGCAG ACGTTTGCGC TCCATCATGG AAAACAGTGG      150
   GTTTGAAGCA TATAGCCTCG AATGGTGGCA CTATGTATTA AGAGACGAAC      200
   CATACCCCAA TAGCTATTTT GATTTCCCCG TTAAATAAAC TTTTAACCGT      250
   TGCACGGACA AACTATATAA GCTAACTCTT TCGGCAGGAA ACCCGACGTA      300
15 TGTAAGTGGT TCTTAGGGAA TTTATATATA GTAGATAGTA TTGAAGATGT      350
   AAGGCAGAGC GATATTGCGG TCATTATCTG CGTGCGCTGC GGCAAGATAG      400
   CCTGATAATA AGACTGATCG CATAGAGGGG TGGTATTTC AACC GCCCAT      450
   TGTCACAGG CAGTTCAGCC TCGTTAAATT CAGCATGGGT ATCACTTATG      500
   AAAATTCATC TACATTGGTG ATAATAGTAA ATCCAGTAGG GCGAAATAAT      550
20 TGACTGTAAT TTACGGGGCA AAACGGCACA ATCTCAAACG AGATTGTGCC      600
   GTTTAAGGGG AAGATTCTAG AAATATTTCA TACTTCCAAC TATATAGTTA      650
   AGGAGGAGAC TGAAAATGAA GAAGTTGTTT TTTTATTGT TATTGTTATT      700
   CTTAATATAC TTAGGTTATT G                                     721

```

25

2) INFORMATION FOR SEQ ID NO: 1068

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 668 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*

(B) STRAIN: R492

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1068

```

45 ATTTTTAAGG ATGAACGCTC TTCATCATGC GGCAAATGGA ATATCATGCA      50
   ATGAAGCGCA AAATCGCAGA CGTTTGCGCT CCATCATGGA AAACAGTGGG      100
   TTTGAAGCAT ATAGCCTCGA ATGGTGGCAC TATGTATTAA GAGACGAACC      150
   ATACCCCAAT AGCTATTTTG ATTTCCCCGT TAAATAAACT TTTAACCGTT      200
   GCACGGACAA ACTATATAAG CTAACCTCTT CGGCAGGAAA CCCGACGTAT      250
   GTAAGTGGTT CTTAGGGAAT TTATATATAG TAGATAGTAT TGAAGATGTA      300
   AGGCAGAGCG ATATTGCGGT CATTATCTGC GTGCGCTGCG GCAAGATAGC      350
50 CTGATAATAA GACTGATCGC ATAGAGGGGT GGTATTTTAC ACCGCCCAT      400
   GTCAACAGGC AGTTCAGCCT CGTTAAATTC AGCATGGGTA TCACTTATGA      450
   AAATTCATCT ACATTGGTGA TAATAGTAAA TCCAGTAGGG CGAAATAAAT      500
   GACTGTAATT TACGGGGCAA AACGGCACA TCTCAAACGA GATTGTGCCG      550
   TTTAAGGGGA AGATTCTAGA AATATTTTCA ACTTCCAAC ATATAGTTAA      600
55 GGAGGAGACT GAAAATGAAG AAGTTGTTTT TTTTATTGTT ATTGTTATTC      650
   TTAATATACT TAGGTTAT                                     668

```

60 2) INFORMATION FOR SEQ ID NO: 1069





ATTTACACACC GCCCATTTGTC AACAGGCAGT TCAGCCTCGT TAAATTCAGC 550  
 ATGGGTATCA CTTATGAAAA TTCATCTACA TTGGTGATAA TAGTAAATCC 600  
 AGTAGGGCGA AATAATTGAC TGTAATTTAC GGGGCAAAAC GGCACAATCT 650  
 CAAACGAGAT TGTGCCGTTT AAGGGGAAGA TTCTAGAAAT ATTTCATACT 700  
 5 TCCAACTATA TAGTTAAGGA GGAGACTGAA AATGAAGAAG TTGTTTTTTT 750  
 TTATTGTTAT TGTTATTCTT AATATACTTA GGTATGACT ACGTTAATGA 800  
 A 801

10

## 2) INFORMATION FOR SEQ ID NO: 1071

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 711 bases  
 15 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 20 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus gallinarum*  
 (B) STRAIN: R684

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1071

TTGTACCAAT GGGGAGCCGA TTTGATTTTA TGGATGAACG CTCTCATCAT 50  
 GCGGCAAATG GAATATCATG CAATGAAGCG CAAAATCGCA GACGTTTGCG 100  
 CTCCATCATG GAAAACAGTG GGTTTGAAGC ATATAGCCTC GAATGGTGCG 150  
 30 ACTATGTATT AAGAGACGAA CCATACCCCA ATAGCTATTT TGATTTCCCC 200  
 GTTAAATAAA CTTTTAACCG TTGCACGGAC AACTATATA AGCTAACTCT 250  
 TTCGGCAGGA AACCCGACGT ATGTAACCTG TTCTTAGGGA ATTTATATAT 300  
 AGTAGATAGT ATTGAAGATG TAAGGCAGAG CGATATTGCG GTCATTATCT 350  
 GCGTGCGCTG CCGCAAGATA GCCTGATAAT AAGACTGATC GCATAGAGGG 400  
 35 GTGGTATTTT ACACCGCCCA TTGTCAACAG GCAGTTCAGC CTCGTTAAAT 450  
 TCAGCATGGG TATCACTTAT GAAAATTCAT CTACATTGGT GATAATAGTA 500  
 AATCCAGTAG GCGGAAATAA TTGACTGTAA TTTACGGGGC AAAACGGCAC 550  
 AATCTCAAAC GAGATTGTGC CGTTTAAGGG GAAGATTCTA GAAATATTTC 600  
 ATACTTCCAA CTATATAGTT AAGGAGGAGA CTGAAAATGA AGAAGTTGTT 650  
 40 TTTTTTATTG TTATTGTTAT TCTTAATATA CTTAGGTTAT GACTACGTTA 700  
 ATGAAGCACT G 711

## 45 2) INFORMATION FOR SEQ ID NO: 1072

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 751 bases  
 50 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 55 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: R688

## 60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1072

```

GCCATTGATC TTACGCTTTA TCGATTAGAC ACGGGTNAGC TTGTACCAAT      50
GGGGAGCCGA TTTGATTTTA TGGATGAACG CTCTCATCAT GCGGCAAATG      100
GAATATCATG CAATGAAGCG CAAAATCGCA GACGTTTGCG CTCCATCATG      150
GAAAACAGTG GGTTTGAAGC ATATAGCCTC GAATGGTGGC ACTATGTATT      200
5  AAGAGACGAA CCATACCCCA ATAGCTATTT TGATTTCCCC GTTAAATAAA      250
CTTTTAACCG TTGCACGGAC AAACATATA AGCTAACTCT TTCGGCAGGA      300
AACCCGACGT ATGTAACCTG TTCTTAGGGA ATTTATATAT AGTAGATAGT      350
ATTGAAGATG TAAGGCAGAG CGATATTGCG GTCATTATCT GCGTGCGCTG      400
CGGCAAGATA GCCTGATAAT AAGACTGATC GCATAGAGGG GTGGTATTTT      450
10 ACACCGCCCA TTGTCAACAG GCAGTTCAGC CTCGTAAAT TCAGCATGGG      500
TATCACTTAT GAAAATTCAT CTACATTGGT GATAATAGTA AATCCAGTAG      550
GGCGAAATAA TTGACTGTAA TTTACGGGGC AAAACGGCAC AATCTCAAAC      600
GAGATTGTGC CGTTTAAGGG GAAGATTCTA GAAATATTTT AACTTCCAA      650
CTATATAGTT AAGGAGGAGA CTGAAAATGA AGAAGTTGTT TTTTATTATT      700
15 TTATTGTTAT TCTTAATATA CTTAGGTTAT GACTACGTTA ATGAAGCACT      750
G

```

## 20 2) INFORMATION FOR SEQ ID NO: 1073

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 685 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 30 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus flavescens*  
 (B) STRAIN: R689

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1073

```

35 ATTTGATTTT ATGGATGAAC GCTCTCATCA TGCGGCAAAT GGAATATCAT      50
GCAATGAAGC GCAAAATCGC AGACGTTTGC GCTCCATCAT GGAAAACAGT      100
GGGTTTGAAG CATATAGCCT CGAATGGTGG CACTATGTAT TAAGAGACGA      150
ACCATACCCC AATAGCTATT TTGATTTCCC CGTTAAATAA ACTTTTAACC      200
40 GTTGCACGGA CAACTATAT AAGCTAACTC TTTCGGCAGG AAACCCGACG      250
TATGTAACCTG GTTCTTAGGG AATTTATATA TAGTAGATAG TATTGAAGAT      300
GTAAGGCAGA GCGATATTGC GGTCATTATC TGCGTGCGCT GCGGCAAGAT      350
AGCCTGATAA TAAGACTGAT CGCATAGAGG GGTGGTATTT CACACCGCCC      400
ATTGTCAACA GGCAGTTCAG CCTCGTTAAA TTCAGCATGG GTATCACTTA      450
45 TGAAAATTCA TCTACATTGG TGATAATAGT AAATCCAGTA GGGCGAAATA      500
ATTGACTGTA ATTTACGGGG CAAAACGGCA CAATCTCAA CGAGATTGTG      550
CCGTTTAAGG GGAAGATTCT AGAAATATTT CATACTTCCA ACTATATAGT      600
TAAGGAGGAG ACTGAAAATG AAGAAGTTGT TTTTATTATT GTTATTGTGA      650
50 TTCTTAATAT ACTTAGGTTA TGACTACGTT AATGA
685

```

## 2) INFORMATION FOR SEQ ID NO: 1074

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 732 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear
- 60

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus faecium*

(B) STRAIN: R690

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1074

```

10 ATCGATTAGA CACGGGTGAG CTTGTACCAA TGGGGAGCCG ATTTGATTTT 50
   ATGGATGAAC GCTCTCATCA TGCGGCAAAT GGAATATCAT GCAATGAAGC 100
   GCAAAATCGC AGACGTTTGC GCTCCATCAT GGAAAACAGT GGGTTTGAAG 150
   CATATAGCCT CGAATGGTGG CACTATGTAT TAAGAGACGA ACCATACCCC 200
   AATAGCTATT TTGATTTCCC CGTTAAATAA ACTTTTAACC GTTGCACGGA 250
   CAACTATAT AAGCTAACTC TTTCGGCAGG AAACCCGACG TATGTAAGTG 300
15 GTTCTTAGGG AATTTATATA TAGTAGATAG TATTGAAGAT GTAAGGCAGA 350
   GCGATATTGC GGTCATTATC TGCCTGCGCT GCGGCAAGAT AGCCTGATAA 400
   TAAGACTGAT CGCATAGAGG GGTGGTATTT CACACCGCCC ATTGTCAACA 450
   GGCAGTTCAG CCTCGTTAAA TTCAGCATGG GTATCACTTA TGAAAATTCA 500
   TCTACATTGG TGATAATAGT AAATCCAGTA GGGCGAAATA ATTGACTGTA 550
20 ATTTACGGGG CAAAACGGCA CAATCTCAAA CGAGATTGTG CCGTTTAAGG 600
   GGAAGATTCT AGAAATATTT CATACTTCCA ACTATATAGT TAAGGAGGAG 650
   ACTGAAAATG AAGAAGTTGT TTTTTTTATT GTTATTGTTA TTCTTAATAT 700
   ACTTAGGTTA TGACTACGTT AATGAAGCAC TG 732

```

25

2) INFORMATION FOR SEQ ID NO: 1075

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 670 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Enterococcus gallinarum*

(B) STRAIN: R691

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1075

```

45 TCTCATCATG CGGCAAATGG AATATCATGC AATGAAGCGC AAAATCGCAG 50
   ACGTTTGCGC TCCATCATGG AAAACAGTGG GTTTGAAGCA TATAGCCTCG 100
   AATGGTGGCA CTATGTATTA AGAGACGAAC CATACCCCAA TAGCTATTTT 150
   GATTCCCCG TTAAATAAAC TTTTAACCGT TGCACGGACA AACTATATAA 200
   GCTAACTCTT TCGGCAGGAA ACCCGACGTA TGTAAGTGGT TCTTAGGGAA 250
   TTTATATATA GTAGATAGTA TTGAAGATGT AAGGCAGAGC GATATTGCGG 300
   TCATTATCTG CGTGCGCTGC GGCAAGATAG CCTGATAATA AGACTGATCG 350
50 CATAGAGGGG TGGTATTTCA CACCGCCCAT TGTCAACAGG CAGTTCAGCC 400
   TCGTTAAATT CAGCATGGGT ATCACTTATG AAAATTTCATC TACATTGGTG 450
   ATAATAGTAA ATCCAGTAGG GCGAAATAAT TGACTGTAAT TTACGGGGCA 500
   AAACGGCACA ATCTCAAACG AGATTGTGCC GTTTAAGGGG AAGATTCTAG 550
   AAATATTTCA TACTTCCAAC TATATAGTTA AGGAGGAGAC TGAAAATGAA 600
55 GAAGTTGTTT TTTTATTGT TATTGTTATT CTTAATATAC TTAGGTTATG 650
   ACTACGTTAA TGAAGCACTG 670

```

60 2) INFORMATION FOR SEQ ID NO: 1076

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: DG131/3 serotype OX3:H8  
 (C) ACCESSION NUMBER: extracted from Z36901

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1076

```

TTGTGTAATA TGAAAATAAT AATTTTTAGA GTGCTAACTT TTTTCTTTGT      50
TATCTTTTCT GTTAATGTGG TTGCGAAGGA ATTTACCTTA GATTTCCTCGA    100
CAGCAAAGAC GTATGTAGAT TCGCTGAATG TCATTGCTC TGCAATAGGT      150
ACTCCATTAC AGACTATTTT ATCAGGAGGT ACGTCTTTAC TGATGATTGA      200
TAGTGGCACA GGGGATAAAT TGTTTGCAGT TGATGTCAGA GGGATAGATC      250
CAGAGGAAGG GCGGTTTAAT AATCTACGGC TTATTGTTGA ACGAAATAAT      300
TTATATGTGA CAGGATTTGT TAACAGGACA AATAATGTTT TTTATCGCTT      350
TGCTGATTTT TCACATGTTA CCTTTCCTGG TACAACGCG GTTACATTGT      400
CTGGTGACAG TAGCTATACC ACGTTACAGC GTGTTGCGGG GATCAGTCGT      450
ACGGGGATGC AGATAAATCG CCATTGCTTG ACTACTTCTT ATCTGGATTT      500
AATGTCGCAT AGCGGAACCT CACTGACGCA GTCTGTGGCA AGAGCGATGT      550
TACGGTTTGT TACTGTGACA GCTGAAGCTT TACGTTTTCG GCAAATTCAG      600
AGGGGATTTT GTACAACACT TGATGATCTC AGTGGGCGTT CTTATGTAAT      650
GACTGCTGAA GATGTTGATC TTACGTTGAA CTGGGGAAGG TTGAGTAGTG      700
TCCTGCCTGA CTATCATGGA CAAGACTCTG TTCGTGTTGG AAGAATTTCT      750
TTTGGGAAGT TTAATGCAAT TCTGGGTAGC GTGGCATTAA TACTGAATTG      800
TCATCATCAT GCATCGCGAG TTGCCAGAAT TGTACCTAAT GAGTTTCCTT      850
CTATGTGCCC GGTAGATGGA AGAGTGCGTG GGATTACGCA CAATAAAATA      900
TTGTGGGACT CATCCACTCT GGGGGCAATT TTGATACGCA GGGCTATTAG      950
CAGTTGAGGG GGGTAAAATG AAAAAAATAT TATTAATAGC TGCATCACTT    1000
TCATTTTTTT CAGCAAGTGT GCTGGCTGCG CCAGATTGTG TAACTGGGAA    1050
GGTGGAGTAT ACAAATATA ATGATGACGA TACCTTTACA GTTAAAGTGG    1100
GAGATAAAGA ATTATTTACT AACAGATGGA ATCTTCAGTC TCTTCTTCTC    1150
AGTGCACAAA TTACGGGGAT GACGGTAACC ATTAAACTA ATGCCTGTCA    1200
TAATGGAGGG GGATTCAGCG AGGTTATTTT CCGTTGACTC AGAATA      1246

```

## 2) INFORMATION FOR SEQ ID NO: 1077

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1259 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (B) STRAIN: 94C serotype O48:H21  
 (C) ACCESSION NUMBER: extracted from Z37725

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1077

CACCTGTATA TGAAGTGTAT ATTATTTAAA TGGGTACTGT GCCTGTACT 50  
 GGGTTTTTCT TCGGTATCCT ATTCCCGGA GTTTACGATA GACTTTTCGA 100  
 CCCAACAAAG TTATGTCTCT TCGTTAAATA GTATACGGAC AGAGATATCG 150  
 5 ACCCCTCTTG AACATATATC TCAGGGGACC ACATCGGTGT CTGTTATTAA 200  
 CCACACCCCA CCGGGCAGTT ATTTTGCTGT GGATATACGA GGGCTTGATG 250  
 TCTATCAGGC GCGTTTTGAC CATCTTCGGC TGATTATTGA GCAAAATAAT 300  
 TTATATGTGG CCGGGTTCGT TAATACGGCA ACAAATACTT TCTACCGTTT 350  
 TTCAGATTTT ACACATATAT CAGTGCCCGA TGTGACAACG GTTTCCATGA 400  
 10 CAACGGACAG CAGTTATACC ACTCTGCAAC GTGTGCGAGC GCTGGAACGT 450  
 TCCGGAATGC AAATCAGTCG TCACTCACTG GTTTCATCAT ATCTGGCGTT 500  
 AATGGAGTTC AGTGGTAATA CAATGACCAG AGATGCATCC AGAGCAGTTC 550  
 TGC GTTTTTGT CACTGTCACA GCAGAAGCCT TACGCTTCAG GCAGATACAG 600  
 AGAGAATTTT GTCAGGCACT GTCTGAAACT GCTCCTGTGT ATACCATGAC 650  
 15 GCCGGGAGAC GTGGACCTCA CTCTGAACTG GGGGCGAATC AGCAATGTGC 700  
 TTCCGGAGTA TCGGGGAGAG GATGGTGTCA GAGTGGGGAG AATATCCTTT 750  
 AATAATATAT CAGCGATACT GGGTACTGTG GCCGTTATAC TGAATTGCCA 800  
 TCATCAGGGG GCGCGTTCG TTCGCGCCGT GAATGAAGAG AGTCAACCAG 850  
 AATGTCAGAT AACTGGCGAC AGGCCCGTTA TAAAAATAAA CAATACATTA 900  
 20 TGGGAAAGTA ATACAGCTGC AGCGTTTCTG AACAGAAAGT CACAGTTTTT 950  
 ATATACAACG GGTAATAAAA GGAGTTAAGT ATGAAGAAGA TGTTTATGGC 1000  
 GGTTTTATTT GCATTAGTTT CTGTAAATGC AATGGCGGCG GATTGTGCTA 1050  
 AAGGTAAAT TGAGTTTTC AAGTATAATG AGGATGACAC ATTTACAGTG 1100  
 AAGGTTGACG GGAAAGAATA CTGGACCAGT CGCTGGAATC TGCAACCGTT 1150  
 25 ACTGCAAAGT GCTCAGCTGA CAGGAATGAC TGTCACAATC AAATCCAGTA 1200  
 CCTGTGAATC AGGCTCCGGA TTTGCTGAAG TGCAGTTTAA TAATGACTGA 1250  
 GGCATAACC 1259

30

## 2) INFORMATION FOR SEQ ID NO: 1078

## (i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1078

AGTTCTGCGT TTTGTCACTG TC

22

45

## 2) INFORMATION FOR SEQ ID NO: 1079

## (i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1079

CGGAAGCACA TTGCTGATT

60

19

## 2) INFORMATION FOR SEQ ID NO: 1080

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 25 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1080

15 TATAGCTACT GTCACCAGAC AATGT 25

## 2) INFORMATION FOR SEQ ID NO: 1081

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1081

30 ATGTCAGAGG GATAGATCCA 20

## 2) INFORMATION FOR SEQ ID NO: 1082

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1082

TTGARCRAAA TAATTTATAT GTG 23

## 2) INFORMATION FOR SEQ ID NO: 1083

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 20 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1083

TGATGATGRC AATTCAGTAT

20

5

2) INFORMATION FOR SEQ ID NO: 1084

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

10

15

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1084

CCACGCCGCT TTGCTGATTT TTCACATGTT ACCGCGTGG

39

20

2) INFORMATION FOR SEQ ID NO: 1085

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

25

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1085

CCACGCCACT GTCTGAAACT GCTCCTGTGC GTGG

34

35

2) INFORMATION FOR SEQ ID NO: 1086

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

40

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1086

CTACTCCCGC CTTTGGGTT

20

50

55 2) INFORMATION FOR SEQ ID NO: 1087

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

60



(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1087

CTCACAGCCC GAAACAGCCT

20

10

2) INFORMATION FOR SEQ ID NO: 1088

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

15 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1088

TGCCGTTTCC TGTATCCGTC

20

25

2) INFORMATION FOR SEQ ID NO: 1089

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1089

ATCCACACGG GCTAGACCTC

20

40

2) INFORMATION FOR SEQ ID NO: 1090

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 21 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1090

55 AATAGCGCGG ACGAATTGGA C

21

55

2) INFORMATION FOR SEQ ID NO: 1091

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1091

AACGCGGCAC TGTTTCCCAA

20

## 2) INFORMATION FOR SEQ ID NO: 1092

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1092

TCGGCAAGAC AATATGACAG C

21

## 2) INFORMATION FOR SEQ ID NO: 1093

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 381 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus saprophyticus*  
 (B) STRAIN: CSsa-165

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1093

TAACGGGCGT	CTCGATAGAA	AAACACGTGA	AAATCCCAAT	GATTATAAAC	50
AATCAATATA	CGATTTTGCT	GAAGCTGTAA	CAAAAGGTAT	TAAGGAACAA	100
ACAAATAAAA	ATTAATAGGC	AACTTAACCA	GAATCGTTAA	AACTATATGA	150
CGATTCTGGT	TTTTTAAATT	CAAAAAGTTT	TCTAAAAAAT	TTACCTGCTT	200
TTTTAAAGTA	TAGGTATAAA	ATACAATTGA	TTAAAATAGT	AAAGGAAATG	250
AATCATGAAA	CAATTAACCTA	AGCCTTTATA	CTTTTACCTA	TTACTTTTTTA	300
TTACAACAAC	ACTGATTGGC	GCGTTACTAT	TATATTTGCC	AATCACAGGT	350
AAACATCCTA	TTGATTTTGT	GGACGCCCGT	T		381

## 2) INFORMATION FOR SEQ ID NO: 1094

## (i) SEQUENCE CHARACTERISTICS:

819

(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1094

10 GGTAACACAG GTACTTCTAA CTA

23

2) INFORMATION FOR SEQ ID NO: 1095

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1095

25

CGATAGAAGC AGCAGGACAA

20

30 2) INFORMATION FOR SEQ ID NO: 1096

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1096

CTGATGGATG CGGAAGATAC

20

45

2) INFORMATION FOR SEQ ID NO: 1097

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1097

TCYTCAAAAG GGATCACWAA AGTMAC

26

60

## 2) INFORMATION FOR SEQ ID NO: 1098

5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

10 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1098

15 TCTTCAAAAT CGAAAAAGCC GTC

23

## 2) INFORMATION FOR SEQ ID NO: 1099

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

25 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1099

30 TCAAAAGGGA TCACWAAAGT MAC

23

## 2) INFORMATION FOR SEQ ID NO: 1100

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 26 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

40 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1100

45 GTAAAKCCCG GCATRGTRTTGATTTC

26

## 50 2) INFORMATION FOR SEQ ID NO: 1101

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1101

GACGGYTTTT TYGATTTTGAAGA

23

5

2) INFORMATION FOR SEQ ID NO: 1102

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1102

AAAAARTCGA TKCGAGCMAGACC

23

20

2) INFORMATION FOR SEQ ID NO: 1103

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1103

ATCCCGCTAT GAAAACGATC

20

35

2) INFORMATION FOR SEQ ID NO: 1104

(i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1104

GGATCAACAC AGTAGAACCG

20

50

2) INFORMATION FOR SEQ ID NO: 1105

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
60 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1105

5

CTCCTACGAT TCTCTTGAYA AATCA

25

10 2) INFORMATION FOR SEQ ID NO: 1106

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

(B) TYPE: Nucleic acid

15

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1106

CAACCGATCT CAACACCGGC AAT

23

25

2) INFORMATION FOR SEQ ID NO: 1107

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

30

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1107

CTCATTTGAC TTCCTCCTTT GCT

23

40

2) INFORMATION FOR SEQ ID NO: 1108

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

45

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1108

GTAAGAATCG GAAAAGCGGA AGG

23

55

2) INFORMATION FOR SEQ ID NO: 1109

(i) SEQUENCE CHARACTERISTICS:

60

823

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1109

10 ACATCGTGAT CGCTAAAAGG AGC

23

## 2) INFORMATION FOR SEQ ID NO: 1110

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1110

25

ACGAGAAAGA CAACAGGAAG ACC

23

## 30 2) INFORMATION FOR SEQ ID NO: 1111

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1111

40

CTTTTCCGG CTCGWYTTCC TGATG

25

## 45 2) INFORMATION FOR SEQ ID NO: 1112

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1112

55

GGCTGYGATA TTCAAAGCTC

20

60

## 2) INFORMATION FOR SEQ ID NO: 1113

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1113

ACCGACCTCA CAGCCCGAAA

20

## 2) INFORMATION FOR SEQ ID NO: 1114

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1114

TCWGAGCCTT TTTCCGGCTC G

21

## 2) INFORMATION FOR SEQ ID NO: 1115

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1115

TTTCGGGCTG TGAGGTCGGB TGHGCG

26

## 2) INFORMATION FOR SEQ ID NO: 1116

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1116

TTTCGGGCTG TGAGGTCGGB TGHGCGG

27

5

2) INFORMATION FOR SEQ ID NO: 1117

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 801 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 20 (A) ORGANISM: *Enterococcus faecium*  
(B) STRAIN: 94  
(C) ACCESSION NUMBER: U94526

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1117

25	AAATTCGATC	CGCACTACAT	CGGAATTACA	AAAAACGGTG	TATGGAAGCT	50
	ATGCAAGAAG	CCATGTACGG	AATGGGAAGC	CGACAGTCTC	CCCGCCATAC	100
	TCTCCCCGGA	TAGGAAAACG	CATGGGCTGC	TTGTCATGAA	AGAAAGCGAA	150
	TACGAAACAC	GGCGTATTGA	TGTGGCTTTC	CCGGTTTTGC	ATGGCAAATG	200
	CGGGGAGGAT	GGTGCATAC	AGGGGCTGTT	TGTATTGTCT	GGTATCCCCT	250
	ATGTGGGCTG	TGATATTCAA	AGCTCCGCAG	CTTGCATGGA	CAAATCACTG	300
30	GCCTACATTC	TTACAAAAAA	TGCGGGCATC	GCCGTTCCCG	AATTTCAAAT	350
	GATTGATAAA	GGTGACAAGC	CGGAGGCGGG	TGCGCTTACC	TACCCTGTCT	400
	TTGTGAAGCC	GGCACGGTCA	GGTTCGTCCT	TTGGCGTAAC	CAAAGTAAAC	450
	GGTACGGAAG	AACTTAACGC	TGCGATAGAA	GCGGCAGGAC	AATATGATGG	500
	AAAAATCTTA	ATTGAGCAAG	CGATTTCCGG	CTGTGAGGTC	GGGTGTGCGG	550
35	TCATGGGGAA	CGAGGATGAT	TTGATTGTCG	GCGAAGTGGA	TCAAATCCGG	600
	CTGAGCCACG	GTATCTTCCG	CATCCATCAG	GAAAACGAGC	CGGAAAAAGG	650
	CTCAGAAAAT	GCGATGATTA	CAGTTCCCGC	AGACATTCCG	GTCGAGGAAC	700
	GAAATCGGGT	GCAGGAAACG	GCAAAGAAAG	TATATCGGGT	GCTTGGATGC	750
	AGAGGGCTTG	CCCCTGTTGA	TCTTTTTTTG	CAGGAGGATG	GCGGCATCGT	800
40	T					801

2) INFORMATION FOR SEQ ID NO: 1118

(i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
50 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1118

55 TTTTCWGAGC CTTTTTCCGG CTCG

24

60 2) INFORMATION FOR SEQ ID NO: 1119

826

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
    (B) TYPE: Nucleic acid  
5      (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1119  
TTTCGGGCTG TGAGGTCGGB TGHGC 25

15 2) INFORMATION FOR SEQ ID NO: 1120

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 24 bases  
20      (B) TYPE: Nucleic acid  
        (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

25   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1120  
TTTCGGGCTG TGAGGTCGGB TGHG 24

30 2) INFORMATION FOR SEQ ID NO: 1121

(i) SEQUENCE CHARACTERISTICS:  
35      (A) LENGTH: 22 bases  
        (B) TYPE: Nucleic acid  
        (C) STRANDEDNESS: Single  
        (D) TOPOLOGY: Linear

40   (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1121  
TGTTTGWATT GTCYGGYATC CC 22

45 2) INFORMATION FOR SEQ ID NO: 1122

50   (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 18 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
55      (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1122

TGGTGCATTG CTACGTGG

18

## 5 2) INFORMATION FOR SEQ ID NO: 1123

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- (B) TYPE: Nucleic acid
- 10 (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1123

TTTCGGGCTG TGAGGTCGGB TG

22

## 20 2) INFORMATION FOR SEQ ID NO: 1124

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases
- 25 (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1124

GATTTGRTCC ACYTCGCCRA CA

22

## 35 2) INFORMATION FOR SEQ ID NO: 1125

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## 45 (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1125

ACTCACAAC TGGGATGGATG

20

50

## 2) INFORMATION FOR SEQ ID NO: 1126

## 55 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

828

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1126

TTATGGTTGT GCTGGTTGAG G

21

2) INFORMATION FOR SEQ ID NO: 1127

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1127

KCAAAYGCCA TTTCAAGTAA

20

2) INFORMATION FOR SEQ ID NO: 1128

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1128

GACGACYTTA TKGATATACA

20

2) INFORMATION FOR SEQ ID NO: 1129

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1129

ATGATGACCG AMATGATGAA AAC

23

2) INFORMATION FOR SEQ ID NO: 1130

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
5      (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1130  
10      GACTATCCAA GCATGCATTA TG 22

15 2) INFORMATION FOR SEQ ID NO: 1131

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 22 bases  
    (B) TYPE: Nucleic acid  
20      (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1131  
    CATCTGGAGC TACRTARCCA GT 22

30 2) INFORMATION FOR SEQ ID NO: 1132

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 20 bases  
35      (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1132  
    AGTGAAAARA TGGCTGCTGC 20

45 2) INFORMATION FOR SEQ ID NO: 1133

(i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 23 bases  
50      (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear

55 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1133  
    CATCAAGAAC ACTGGCTAYG TAG 23

## 2) INFORMATION FOR SEQ ID NO: 1134

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1134

CTAGATAGAG CTAAAACCTT CCT

23

## 2) INFORMATION FOR SEQ ID NO: 1135

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1135

CATTATGCAA ACGCCATTTT AAG

23

## 2) INFORMATION FOR SEQ ID NO: 1136

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1136

ACTTGTCAC GTTSGATRTC T

21

## 2) INFORMATION FOR SEQ ID NO: 1137

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

831

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1137

AATTAATGGC TGCWGTGAY GAA

23

## 2) INFORMATION FOR SEQ ID NO: 1138

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1347 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus gallinarum*
- (C) ACCESSION NUMBER: M75132

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1138

25	GAGCTCTTCC	TTCAACGCAC	TTCTGTACCA	AGAGTTGTTG	TCCATTTGAT	50
	CACTAACAAT	AGCTTCCCCT	GCTTTCTTCA	AGCCCTTTGT	CATAAAATCG	100
	TTAGATTTTC	ATCATAAAAA	TACGAGAAAG	ACAACAGGAA	GACCGCAAAT	150
	TTTCTTTTCT	TTTCCTAGGT	ACACTGAATG	TAACCTTAAA	AGAAAAAAGG	200
	AAAGGAAGAA	AATGATGAAA	AAAATTGCCG	TTTTATTTGG	AGGGAATTCT	250
	CCAGAATACT	CAGTGTCACT	AACCTCAGCA	GCAAGTGTGA	TCCAAGCTAT	300
30	TGACCCGCTG	AAATATGAAG	TAATGACCAT	TGGCATCGCA	CCAACAATGG	350
	ATTGGTATTG	GTATCAAGGA	AACCTCGCGA	ATGTTTCGCA	TGATACTTGG	400
	CTAGAAGATC	ACAAAACTG	TCACCAGCTG	ACTTTTCTTA	GCCAAGGATT	450
	TATATTAGGA	GAAAAACGAA	TCGTCCCTGA	TGTCCTCTTT	CCAGTCTTGC	500
	ATGGGAAGTA	TGGCGAGGAT	GGCTGTATCC	AAGGACTGCT	TGAACATAATG	550
35	AACCTGCCTT	ATGTTGGTTG	CCATGTCGCT	GCCTCCGCAT	TATGTATGAA	600
	CAAATGGCTC	TTGCATCAAC	TTGCTGATAC	CATGGGAATC	GCTAGTGCTC	650
	CCACTTTGCT	TTTATCCCGC	TATGAAAACG	ATCCTGCCAC	AATCGATCGT	700
	TTTATTCAAG	ACCATGGATT	CCCGATCTTT	ATCAAGCCGA	ATGAAGCCGG	750
	TTCTTCAAAA	GGGATCACAA	AAGTAACTGA	CAAAACAGCG	CTCCAATCTG	800
40	CATTAACGAC	TGCTTTTGCT	TACGGTTCTA	CTGTGTTGAT	CCAAAAGGCG	850
	ATAGCGGGTA	TTGAAATTGG	CTGCGGCATC	TTAGGAAATG	AGCAATTGAC	900
	GATTGGTGCT	TGTGATGCGA	TTTCTCTTGT	CGACGGTTTT	TTTGATTTTG	950
	AAGAGAAATA	CCAATTAATC	AGCGCCACGA	TCACTGTCCC	AGCACCATTG	1000
	CCTCTCGCGC	TTGAATCACA	GATCAAGGAG	CAGGCACAGC	TGCTTTATCG	1050
45	AAACTTGGGA	TTGACGGGTC	TGGCTCGAAT	CGATTTTTTC	GTCACCAATC	1100
	AAGGAGCGAT	TTATTTAAAC	GAAATCAACA	CCATGCCGGG	ATTTACTGGG	1150
	CACTCCCGCT	ACCCAGCTAT	GATGGCGGAA	GTCGGGTTAT	CCTACGAAAT	1200
	ATTAGTAGAG	CAATTGATTG	CACTGGCAGA	GGAGGACAAA	CGATGAACAC	1250
	ATTACAATTG	ATCAATAAAA	ACCATCCATT	GAAAAAAAAT	CAAGAGCCCC	1300
50	CGCACTTAGT	GCTAGCTCCT	TTTAGCGATC	ACGATGTTTA	CCTGCAG	1347

## 2) INFORMATION FOR SEQ ID NO: 1139

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1768 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- 5 (A) ORGANISM: *Enterococcus faecium*  
 (B) STRAIN: BM4147  
 (C) ACCESSION NUMBER: X56895

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1139

```

10 GATATCGTTA CGCTTCATGT GCCGCTCAAT ACGGATACGC ACTATATTAT 50
   CAGCCACGAA CAAATACAGA GAATGAAGCA AGGAGCATTT CTTATCAATA 100
   CTGGGCGCGG TCCACTTGTA GATACCTATG AGTTGGTTAA AGCATTAGAA 150
   AACGGGAAAC TGGGCGGTGC CGCATTGGAT GTATTGGAAG GAGAGGAAGA 200
15 GTTTTTCTAC TCTGATTGCA CCCAAAACC AATTGATAAT CAATTTTAC 250
   TTAAACTTCA AAGAATGCCT AACGTGATAA TCACACCGCA TACGGCCTAT 300
   TATACCGAGC AAGCGTTGCG TGATACCGTT GAAAAAACCA TTAAAAACTG 350
   TTTGGATTTT GAAAGGAGAC AGGAGCATGA ATAGAATAAA AGTTGCAATA 400
   CTGTTTGGGG GTTGCTCAGA GGAGCATGAC GTATCGGTAA AATCTGCAAT 450
20 AGAGATAGCC GCTAACATTA ATAAAGAAAA ATACGAGCCG TTATACATTG 500
   GAATTACGAA ATCTGGTGTA TGGAAAATGT GCGAAAAACC TTGCGCGGAA 550
   TGGGAAAACG ACAATTGCTA TTCAGCTGTA CTCTCGCCGG ATAAAAAAT 600
   GCACGGATTA CTTGTAAAAA AGAACCATGA ATATGAAATC AACCATGTTG 650
   ATGTAGCATT TTCAGCTTTG CATGGCAAGT CAGGTGAAGA TGGATCCATA 700
25 CAAGGTCTGT TTGAATTGTC CGGTATCCCT TTTGTAGGCT GCGATATTCA 750
   AAGCTCAGCA ATTTGTATGG ACAAATCGTT GACATACATC GTTGCGAAAA 800
   ATGCTGGGAT AGCTACTCCC GCCTTTTGGG TTATTAATAA AGATGATAGG 850
   CCGGTGGCAG ATACGTTTAC CTATCCTGTT TTTGTTAAGC CGGCGCGTTC 900
   AGGCTCATCC TTCGGTGTGA AAAAAGTCAA TAGCGCGGAC GAATTGGACT 950
30 ACGCAATTGA ATCGGCAAGA CAATATGACA GCAAATCTT AATTGAGCAG 1000
   GCTGTTTCGG GCTGTGAGGT CGGTTGTGCG GTATTGGGAA ACAGTGCCGC 1050
   GTTAGTTGTT GGCGAGGTGG ACCAAATCAG GCTGCAGTAC GGAATCTTTC 1100
   GTATTCATCA GGAAGTCGAG CCGGAAAAAG GCTCTGAAAA CGCAGTTATA 1150
   ACCGTTCCCC CAGACCTTTC AGCAGAGGAG CGAGGACGGA TACAGGAAAC 1200
35 GGCAAAAAAA ATATATAAAG CGCTCGGCTG TAGAGGTCTA GCCCGTGTGG 1250
   ATATGTTTTT ACAAGATAAC GGCCGCATTG TACTGAACGA AGTCAATACT 1300
   CTGCCCCGGT TCACGTCATA CAGTCGTTAT CCCCCTATGA TGGCCGCTGC 1350
   AGGTATTGCA CTTCCCGAAC TGATTGACCG CTTGATCGTA TTAGCGTTAA 1400
   AGGGGTGATA AGCATGGAAA TAGGATTTAC TTTTTTAGAT GAAATAGTAC 1450
40 ACGGTGTTTC TTGGGACGCT AAATATGCCA CTTGGGATAA TTTCACCGGA 1500
   AAACCGGTTG ACGGTTATGA AGTAAATCGC ATTGTAGGGA CATACGAGTT 1550
   GGCTGAATCG CTTTTGAAGG CAAAAGAACT GGCTGCTACC CAAGGGTACG 1600
   GATTGCTTCT ATGGGACGGT TACCGTCCTA AGCGTGCTGT AAAGTTATTA 1650
   ATGCAATGGG CTGCACAGCC GGAAAATAAC CTGACAAAGG AAAGTTATTA 1700
45 TCCCAATATT GACCGAAGT AGATGATTTT AAAAGGATAC GTGGCTTCAA 1750
   AATCAAGCCA TAGCCCGC 1768

```

50 2) INFORMATION FOR SEQ ID NO: 1140

(i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 1086 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

60 (vi) ORIGINAL SOURCE:



(A) ORGANISM: *Enterococcus casseliflavus*  
 (C) ACCESSION NUMBER: L29638

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1140

```

5   GTAAGAATCG GAAAAGCGGA AGGAAGAAAA ACATGAAAAA AATCGCCATT      50
    ATTTTGGGAG GCAATTCACC GGAATACACC GTTCTTTAG CTTCAGCAAC      100
    TAGCGCAATC GAAGCACTCC AATCATCTCC CTATGACTAC GACCTCTCTT      150
    TGATCGGGAT CGCCCCAGAT GCTATGGATT GGTACTTGTA TACAGGAGAA      200
10  CTGGAAAACA TCCGACAAGA CACGTGGTTG TTGGATACGA AACATAAACA      250
    GAAAATACAG CCGCTATTCTG AAGGAAACGG CTTTGGCTA AGTGAAGAGC      300
    AGCAAACGTT GGTACCTGAT GTTTTATTTT CCATTATGCA TGGCAAATAC      350
    GGGGAAGATG GCAGTATCCA AGGATTGTTT GAATTGATGA AGCTGCCTTA      400
    TGTAGGCTGC GGGGTGGCAG GTTCTGCCTT ATGTATGAAC AAATGGCTGC      450
15  TGCATCAAGC TGCAGCAGCC ATTGGCGTAC AAAGTGCTCC TACGATTCTC      500
    TTGACAAATC AAGCCAACCA GCAAGAACAA ATCGAAGCTT TTATCCAGAC      550
    CCATGGCTTC CCAGTTTCTT TTAAGCCTAA TGAAGCGGGC TCCTCAAAG      600
    GGATCACTAA AGTCACCTGC GTTGAAGAAA TCGCTTCTGC CTTAAAAGAA      650
    GCCTTTACTT ATTGTTCCGC AGTGCTCCTA CAAAAAATA TTGCCGGTGT      700
20  TGAGATCGGT TCGGGTATTT TGGGCAACGA CTCCTTGACT GTCGGTGCTT      750
    GTGACGCCAT TTCATTAGTA GACGGCTTTT TCGATTTTGA AGAAAAGTAC      800
    CAGCTGATCA GCGCCAAAAT CACCGTCCCT GCGCCATTGC CTGAAACGAT      850
    TGAAACCAAG GTCAAAGAAC AAGCTCAGCT GCTCTATCGT AGTCTTGGTC      900
    TTAAAGGTCT TGCTCGCATC GACTTTTTTG TCACGGAGCG AGGAGAACTA      950
25  TACTTGAATG AAATCAATAC TATGCCGGGC TTTACGAGTC ACTCCCGCTA      1000
    TCCTGCCATG ATGGCAGCGG TCGGCTTATC CTATCAAGAA CTACTACAAA      1050
    AACTGCTTGT CTTAGCAAAG GAGGAAGTCA AATGAG                      1086
  
```

## 30 2) INFORMATION FOR SEQ ID NO: 1141

## (i) SEQUENCE CHARACTERISTICS:

```

    (A) LENGTH: 3946 bases
    (B) TYPE: Nucleic acid
35  (C) STRANDEDNESS: Double
    (D) TOPOLOGY: Linear
  
```

## (ii) MOLECULE TYPE: Genomic DNA

## 40 (vi) ORIGINAL SOURCE:

```

    (A) ORGANISM: Enterococcus faecium
    (B) STRAIN: BM4147
    (C) ACCESSION NUMBER: extracted from M97297
  
```

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1141

```

    ATGAATAACA TCGGCATTAC TGTTTATGGA TGTGAGCAGG ATGAGGCAGA      50
    TGCATTCCAT GCTCTTTCGC CTCGCTTTGG CGTTATGGCA ACGATAATTA      100
    ACGCCAACGT GTCGGAATCC AACGCCAAAT CCGCGCCTTT CAATCAATGT      150
50  ATCAGTGTGG GACATAAATC AGAGATTTC GCCTCTATTC TTCTTGCGCT      200
    GAAGAGAGCC GGTGTGAAAT ATATTTCTAC CCGAAGCATC GGCTGCAATC      250
    ATATAGATAC AACTGCTGCT AAGAGAATGG GCATCACTGT CGACAATGTG      300
    GCGTACTCGC CGGATAGCGT TGCCGATTAT ACTATGATGC TAATTCTTAT      350
    TGCAGTACGC AACGTAAAAT CGATTGTGCG CTCTGTGGAA AAACATGATT      400
55  GCGAGTTGGA CAGCGACCGT GGCAAGGTAC TCAGCGACAT GACAGTTGGT      450
    GTGGTGGGAA CGGGCCAGAT AGGCAAAGCG GTTATTGAGC GGCTGCGAGG      500
    ATTTGGATGT AAAGTGTGG CTTATAGTCG CAGCCGAAGT ATAGAGGTAA      550
    ACTATGTACC GTTTGATGAG TTGCTGCAAA ATAGCGATAT CGTTACGCTT      600
    CATGTGCCGC TCAATACGGA TACGCACTAT ATTATCAGCC ACGAACAAAT      650
60  ACAGAGAATG AAGCAAGGAG CATTTCTTAT CAATACTGGG CGCGGTCCAC      700
  
```

	TTGTAGATAC	CTATGAGTTG	GTTAAAGCAT	TAGAAAACGG	GAAACTGGGC	750
	GGTGCCGCAT	TGGATGTATT	GGAAGGAGAG	GAAGAGTTTT	TCTACTCTGA	800
	TTGCACCCAA	AAACCAATTG	ATAATCAATT	TTTACTTAAA	CTTCAAAGAA	850
	TGCCTAACGT	GATAATCACA	CCGCATACGG	CCTATTATAC	CGAGCAAGCG	900
5	TTGCGTGATA	CCGTTGAAAA	AACCATTAAA	AACTGTTTGG	ATTTTGAAAG	950
	GAGACAGGAG	CATGAATAGA	ATAAAAGTTG	CAATACTGTT	TGGGGGTGTC	1000
	TCAGAGGAGC	ATGACGTATC	GGTAAAATCT	GCAATAGAGA	TAGCCGCTAA	1050
	CATTAATAAA	GAAAAATACG	AGCCGTTATA	CATTGGAATT	ACGAAATCTG	1100
	GTGTATGGAA	AATGTGCGAA	AAACCTTGCG	CGGAATGGGA	AAACGACAAT	1150
10	TGCTATTTCAG	CTGTACTCTC	GCCGGATAAA	AAAATGCACG	GATTACTTGT	1200
	TAAAAAGAAC	CATGAATATG	AAATCAACCA	TGTTGATGTA	GCATTTTCAG	1250
	CTTTGCGATG	CAAGTCAGGT	GAAGATGGAT	CCATACAAGG	TCTGTTTGAA	1300
	TTGTCCGGTA	TCCCTTTTGT	AGGCTGCGAT	ATTCAAAGCT	CAGCAATTTG	1350
	TATGGACAAA	TCGTTGACAT	ACATCGTTGC	GAAAAATGCT	GGGATAGCTA	1400
15	CTCCCGCCTT	TTGGGTTATT	AATAAAGATG	ATAGGCCGGT	GGCAGCTACG	1450
	TTTACCTATC	CTGTTTTTGT	TAAGCCGGCG	CGTTCAGGCT	CATCCTTCGG	1500
	TGTGAAAAAA	GTCAATAGCG	CGGACGAATT	GGACTACGCA	ATTGAATCGG	1550
	CAAGACAATA	TGACAGCAAA	ATCTTAATTG	AGCAGGCTGT	TTCGGGCTGT	1600
	GAGGTCGGTT	GTGCGGTATT	GGGAAACAGT	GCCGCGTTAG	TTGTTGGCGA	1650
20	GGTGGACCAA	ATCAGGCTGC	AGTACGGAAT	CTTTCGTATT	CATCAGGAAG	1700
	TCGAGCCGGA	AAAAGGCTCT	GAAAACGCAG	TTATAACCGT	TCCCGCAGAC	1750
	CTTTCAGCAG	AGGAGCGAGG	ACGGATACAG	GAAACGGCAA	AAAAAATATA	1800
	TAAAGCGCTC	GGCTGTAGAG	GTCTAGCCCG	TGTGGATATG	TTTTTACAAG	1850
	ATAACGCCCG	CATTGTACTG	AACGAAGTCA	ATACTCTGCC	CGGTTTCACG	1900
25	TCATACAGTC	GTTATCCCCG	TATGATGGCC	GCTGCAGGTA	TTGCACTTCC	1950
	CGAACTGATT	GACCGCTTGA	TCGTATTAGC	GTTAAAGGGG	TGATAAGCAT	2000
	GGAAATAGGA	TTTACTTTTT	TAGATGAAAT	AGTACACGGT	GTTTCGTTGGG	2050
	ACGCTAAATA	TGCCACTTGG	GATAATTTCA	CCGGAAAACC	GGTTGACGGT	2100
	TATGAAGTAA	ATCGCATTTG	AGGGACATAC	GAGTTGGCTG	AATCGCTTTT	2150
30	GAAGGCAAAA	GAACCTGGCTG	CTACCCAAGG	GTACGGATTG	CTTCTATGGG	2200
	ACGGTTTACC	TCCTAAGCGT	GCTGTAGCGT	GTTTTATGCA	ATGGGCTGCA	2250
	CAGCCGGAAA	ATAACCTGAC	AAAGGAAAGT	TATTATCCCA	ATATTGACCG	2300
	AACTGAGATG	ATTTCAAAG	GATACGTGGC	TTCAAATCA	AGCCATAGCC	2350
	GCGGCAGTGC	CATTGATCTT	ACGCTTTTATC	GATTAGACAC	GGGTGAGCTT	2400
35	GTACCAATGG	GGAGCCGATT	TGATTTTATG	GATGAACGCT	CTCATCATGC	2450
	GGCAAATGGA	ATATCATGCA	ATGAAGCGCA	AAATCGCAGA	CGTTTGCGCT	2500
	CCATCATGGA	AAACAGTGGG	TTTGAAGCAT	ATAGCCTCGA	ATGGTGGCAC	2550
	TATGTATTAA	GAGACGAACC	ATACCCCAAT	AGCTATTTTG	ATTTCCCCGT	2600
	TAAATAAACT	TTTAACCGTT	GCACGGACAA	ACTATATAAG	CTAACTCTTT	2650
40	CGGCAGGAAA	CCCGACGTAT	GTAACCTGGT	CTTAGGGAAT	TTATATATAG	2700
	TAGATAGTAT	TGAAGATGTA	AGGCAGAGCG	ATATTGCGGT	CATTATCTGC	2750
	GTGCGCTGCG	GCAAGATAGC	CTGATAATAA	GACTGATCGC	ATAGAGGGGT	2800
	GGTATTTTAC	ACCGCCCATT	GTCAACAGGC	AGTTCAGCCT	CGTTAAATTC	2850
	AGCATGGGTA	TCACTTATGA	AAATTTCATCT	ACATTGGTGA	TAATAGTAAA	2900
45	TCCAGTAGGG	CGAAATAATT	GACTGTAATT	TACGGGGCAA	AACGGCACAA	2950
	TCTCAAACGA	GATTGTGCCG	TTTAAGGGGA	AGATTCTAGA	AATATTTTCAT	3000
	ACTTCCAAC	ATATAGTTAA	GGAGGAGACT	GAAAATGAAG	AAGTTGTTTT	3050
	TTTTATTGTT	ATTGTTATTC	TTAATATACT	TAGGTTATGA	CTACGTTAAT	3100
	GAAGCACTGT	TTTCTCAGGA	AAAAGTCGAA	TTTCAAAATT	ATGATCAAAA	3150
50	TCCCAAAGAA	CATTTAGAAA	ATAGTGGGAC	TTCTGAAAAT	ACCCAAGAGA	3200
	AAACAATTAC	AGAAGAACAG	GTTTATCAAG	GAAATCTGCT	ATTAATCAAT	3250
	AGTAAATATC	CTGTTGCGCA	AGAAAGTGTG	AAGTCAGATA	TCGTGAATTT	3300
	ATCTAAACAT	GACGAATTAA	TAAATGGATA	CGGGTTGCTT	GATAGTAATA	3350
	TTTATATGTC	AAAAGAAATA	GCACAAAAAT	TTTCAGAGAT	GGTCAATGAT	3400
55	GCTGTAAAGG	GTGGCGTTAG	TCATTTTATT	ATTAATAGTG	GCTATCGAGA	3450
	CTTTGATGAG	CAAAGTGTGC	TTTACCAAGA	AATGGGGGCT	GAGTATGCCT	3500
	TACCAGCAGG	TTATAGTGAG	CATAATTTCAG	GTTTATCACT	AGATGTAGGA	3550
	TCAAGCTTGA	CGAAAATGGA	ACGAGCCCCT	GAAGGAAAGT	GGATAGAAGA	3600
	AAATGCTTGG	AAATACGGGT	TCATTTTACC	TTATCCAGAG	GACAAAACAG	3650
60	AGTTAACAGG	AATTCAATAT	GAACCATGGC	ATATTGCTA	TGTTGGTTTA	3700

	CCACATAGTG	CGATTATGAA	AGAAAAGAAT	TTCGTTCTCG	AGGAATATAT	3750
	GGATTACCTA	AAAGAAGAAA	AAACCATTTT	TGTTAGTGTA	AATGGGGAAA	3800
	AATATGAGAT	CTTTTATTAT	CCTGTTACTA	AAAATACCAC	CATTCATGTG	3850
	CCGACTAATC	TTCGTTATGA	GATATCAGGA	AACAATATAG	ACGGTGTAAT	3900
5	TGTGACAGTG	TTTCCCGGAT	CAACACATAC	TAATTCAAGG	AGGTAA	3946

## 2) INFORMATION FOR SEQ ID NO: 1142

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1142

GATCCTCTAA ATGATTCTCA GGTGG

25

## 2) INFORMATION FOR SEQ ID NO: 1143

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1143

CAATTAGCTT AGCAATAGGT GTTGG

25

## 2) INFORMATION FOR SEQ ID NO: 1144

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1144

TGTYTTCCAA GGTTCAGCTC

20

## 2) INFORMATION FOR SEQ ID NO: 1145

## (i) SEQUENCE CHARACTERISTICS:

836

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

5

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1145

10 AACATATTKG GTTGATAGGT

20

## 2) INFORMATION FOR SEQ ID NO: 1146

15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1146

25

GGGATTACCT ATGCCAATATGAT

23

## 30 2) INFORMATION FOR SEQ ID NO: 1147

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1147

40

AGCTGTGTTA GCVCGAACAT CTTG

24

45

## 2) INFORMATION FOR SEQ ID NO: 1148

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1148

55

GACTTTGT TT GCGTGATAT

20

## 2) INFORMATION FOR SEQ ID NO: 1149

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1149  
15 TCCYACWATT TCTTTTTGWWG

20

## 2) INFORMATION FOR SEQ ID NO: 1150

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1150  
30 TGATAATCAC ACCGCATACG

20

## 35 2) INFORMATION FOR SEQ ID NO: 1151

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1151  
TGCTGTCATA TTGTCTTGCC

20

## 50 2) INFORMATION FOR SEQ ID NO: 1152

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
55 (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1152

ATAAAGATGA TAGGCCGGTG

20

2) INFORMATION FOR SEQ ID NO: 1153

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1153

CTCGTATGTC CCTACAATGC

20

2) INFORMATION FOR SEQ ID NO: 1154

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1154

GTTTGAAGCA TATAGCCTCG

20

2) INFORMATION FOR SEQ ID NO: 1155

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1155

CAGTGCTTCA TTAACGTAGT C

21

2) INFORMATION FOR SEQ ID NO: 1156

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases

- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1156

GTTGAAATGC ATCACGAACA ATT

23

10

2) INFORMATION FOR SEQ ID NO: 1157

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1157

25 AAGAACGTTT CAGTTAAGGA AAT

23

2) INFORMATION FOR SEQ ID NO: 1158

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1158

40

AAGAGGTAAT GTCTGTGGT

19

45 2) INFORMATION FOR SEQ ID NO: 1159

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1159

TGAAGGTTTG CCAGGTGA

18

## 2) INFORMATION FOR SEQ ID NO: 1160

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1160

CGTTTCTGTT AAAGAAATTA GAAG

24

## 2) INFORMATION FOR SEQ ID NO: 1161

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1161

TCCAGGTGAT AACGTTGG

18

## 2) INFORMATION FOR SEQ ID NO: 1162

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1162

CAAGTCCGTG GAAATGCA

18

## 2) INFORMATION FOR SEQ ID NO: 1163

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1163

GTTGGTTTCA ACGTTAAGAA C

21

5

2) INFORMATION FOR SEQ ID NO: 1164

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 18 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1164

GGTTTCAACG TCAAGAAC

18

20

2) INFORMATION FOR SEQ ID NO: 1165

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

30 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1165

35 GGTTTCAACG TGAAGAAC

18

2) INFORMATION FOR SEQ ID NO: 1166

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1166

50

ACGTTAAGAA TGTTTCTGTC AA

22

55 2) INFORMATION FOR SEQ ID NO: 1167

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 bases  
(B) TYPE: Nucleic acid

842

(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1167

GTTGGTTTCA ACGT

14

10

2) INFORMATION FOR SEQ ID NO: 1168

(i) SEQUENCE CHARACTERISTICS:

15

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1168

GAACAATTGG TTGAAGGTGT

20

25

2) INFORMATION FOR SEQ ID NO: 1169

(i) SEQUENCE CHARACTERISTICS:

30

(A) LENGTH: 2160 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

40

(A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: SP-665  
(C) ACCESSION NUMBER: AF139883

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1169

45	ATGAACAAAC	CAACGATTCT	GCGCCTAATC	AAGTATCTGA	GCATTAGCTT	50
	CTTAAGCTTG	GTTATCGCAG	CCATTGTCTT	AGGCGGAGGA	GTTTTTTTCT	100
	ACTACGTTAG	CAAGGCTCCT	AGCCTATCCG	AGAGTAAACT	AGTTGCAACA	150
	ACTTCTAGTA	AAATCTACGA	CAATAAAAAT	CAACTCATTG	CTGACTTGGG	200
	TTCTGAACGC	CGCGTCAATG	CCCAAGCTAA	TGATATTCCC	ACAGATTTGG	250
50	TTAAGGCAAT	CGTTTCTATC	GAAGACCATC	GCTTCTTCGA	CCACAGGGGG	300
	ATTGATTCCA	TCCGTATCCT	GGGAGCTTTC	TTGCGCAATC	TGCAAAGTAA	350
	TTCCCTCCAA	GGTGGATCAA	CTCTCACCCA	ACAGTTGATT	AAGTTGACTT	400
	ACTTTTCAAC	CTCGACTTCC	GACCAGACTA	TTTCTCGTAA	GGCTCAGGAA	450
	GCTTGGTAG	CGATTCAAGT	AGAACAAAAA	GCAACCAAAC	AGGAAATCTT	500
55	GACCTACTAT	ATAAATAAGG	TCTACATGTC	TAATGGCAAC	TATGGAATGC	550
	AGACAGCAGC	TCAAACTAC	TATGGTAAAG	ACCTCAATAA	TTTAAGTTTA	600
	CCTCAGTTAG	CCTTGCTGGC	TGGAATGCCT	CAGGCACCAA	ACCAATATGA	650
	CCCTATTCA	CATCCAGAAG	CAGCCCAAGA	CCGCCGAAAC	TTGGTCTTAT	700
	CTGAAATGAA	AAATCAAGGT	TACATCTCTG	CTGAACAGTA	TGAGAAAGCA	750
60	GTCAATACAC	CAATTACTGA	TGGACTACAA	AGTCTCAAAT	CAGCAAGTAA	800

```

5  TTACCCTGCT TACATGGATA ATTACCTCAA GGAGGTCATC AATCAAGTAG 850
   AACAAGAAAC TGGCTATAAC CTTCTAACTA CTGGGATGGA TGTTTACACA 900
   AATGTAGACC AAGAAGCTCA AAAACATCTG TGGGATATCT ACAACTCCGA 950
   TCAATACGTC TCTTACCCTG ACGATGATTT GCAAGTCGCA TCTACGGTCG 1000
   TAGATGTTTC AAATGGTAAA GTCATCGCCC AACTTGGAGC TCGTCACCAA 1050
   GCAAGTAACG TTTCATTTGG TACCAACCAA GCTGTGGAAA CCAATCGTGA 1100
   CTGGGGTTCT GCTATGAAAC CAATCACCGA TTATGCACCT GCCATAGAAT 1150
   ACGGTGTTTA TGATTCCACT GCAACTATGG TTAATGATAT TCCTTATAAC 1200
   TATCCGGGAA CAAGCACACC TGTCTACAAC TGGGATAGAG CATATTTTCGG 1250
10 TAATATTACT CTGCAATATG CTCTTCAACA ATCACGAAAT GTCACAGCCG 1300
   TTGAGACTTT GAATAAGGTC GGTCTAGATA GAGCTAAAAC CTTCTTAAAT 1350
   GGTCTTGGTA TCGACTATCC AAGCATGCAT TATGCAAACG CCATTTCAAG 1400
   TAATACAACA GAATCTAATA AACAATACGG AGCAAGTAGT GAAAAAATGG 1450
   CTGCTGCTTA TGCTGCCTTT GCAAATGGTG GCACTTACTA TAAACCAATG 1500
15 TATATCCATA AAGTCGTCTT CAGTGATGGA AGTAAAAAAG AGTTCTCTAA 1550
   TGTCGGAACG CGTGCCATGA AGGAAACGAC TGGGATAGAG ATGACCGACA 1600
   TGATGAAAAC AGTCTTGACT TATGGAACG GCGGTGGAGC CTATCTTCCT 1650
   TGGCTTCCCTC AAGCTGGTAA AACAGGAACC TCTAACTATA CAGATGAGGA 1700
   AGTTGAAAAC CACATCAAGA ACACTGGCTA TGTAGCTCCA GATGAAATGT 1750
20 TTGTTGGTTA TACTCGTAAG TATTCTATGG CTGTATGGAC AGGTTATTCG 1800
   AATCGTTTAA CTCCTATCGT TGGAGATGGT TTCCTAGTTG CAGCTAAAGT 1850
   TTATCGCTCA ATGATAACGT ATCTATCAGA AGATACTCAT CCAGAAGACT 1900
   GGACGATGCC AGACGGACTT TTCAGAAACG GGAATTTTGT ATTCAAAAAT 1950
   GGAGCTCGCC CAATATGGAC TGAACCCCTCT ACTCAACAAT CCTCAACAGC 2000
25 TGAAAGTTCA AGCTCATCAT CAGATAGTTC AACTTCACAG TCTAGCTCAA 2050
   CCACTCCAAG CACAAATAAT AGTACGACTA CCAATCCTAA CAATAATACG 2100
   CAACAATCAA ATACAACCCC TGATCAACAA AATCAGAATC CTCAACCAGC 2150
   ACAACCATAA 2160

```

30

## 2) INFORMATION FOR SEQ ID NO: 1170

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1170

ACGAATTGGA CTACGCAATT

20

45

## 2) INFORMATION FOR SEQ ID NO: 1171

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: DNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1171

ACGAGGATGA TTTGATTGTC

20

60

## 2) INFORMATION FOR SEQ ID NO: 1172

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1560 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: 64147  
 (C) ACCESSION NUMBER: X13136

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1172

20	GATCCTCTAA	ATGATTCTCA	GGTGGCTGTT	ATTGCCTCTA	TTTCAAAGGA	50
	GATGCCTGGC	ATTAGTATTT	CTACTTCTTG	GGATAGAAAG	GTTTGGGAAA	100
	CTTCCCTTTC	TTCTATAGTT	GGGAGTGTAT	CCAGTGAAAA	AGCTGGTCTC	150
	CCAGCGGAAG	AAGCAGAAGC	CTATCTTAAA	AAAGGCTATT	CTCTAAATGA	200
25	CCGTGTAGGA	ACCTCCTATT	TGGAAAAGCA	ATATGAAGAG	ACCTTACAAG	250
	GAAAACGCTC	GGTAAAAGAA	ATCCATCTGG	ACAAATATGG	CAACATGGAA	300
	AGCGTGGATA	CAATTGAGGA	AGGTAGTAAG	GGAAACAATA	TCAAGCTGAC	350
	CATTGATTTG	GCCTTCCAAG	ATAGCGTGGA	TGCTTTGCTG	AAAAGTTATT	400
	TCAATTCAGA	GTTGGGAAAT	GGTGGAGCCA	AGTATTCTGA	AGGTGTCTAT	450
30	GCAGTCGCCC	TTAACCCAAA	AACAGGTGCT	GTTTTGTCTA	TGTCAGGAAT	500
	TAAACATGAC	TTGAAAACAG	GAGAGTTGAC	GCCGGATTCC	TTGGGAACGG	550
	TAACCAATGT	CTTTGTCCCA	GGTTCGGTTG	TCAAAGCAGC	GACCATCAGC	600
	TCTGGTTGGG	AAAATGGAGT	CTTGTCAGGA	AATCAGACCT	TGACAGACCA	650
	GTCCATTGTC	TTTCAAGGTT	CAGCTCCAAT	TAATTCTTGG	TATCCTGCCT	700
35	TTTCTAGACC	AATGCCGATT	ACGGCGGTTT	AGGCTCTAGA	GTATTCATCC	750
	AATGCTTATA	TGGTCCAAAC	AGCCCTAGGT	CTTATGGGGC	AGACCTATCA	800
	ACCCAATATG	TTTGTCGGCA	CCAGCAATCT	AGAGTCTGCT	ATGGGGAAAT	850
	TGCGTTCAAC	CTTTGGTGAA	TATGGTTTGG	GTTCTGCGAC	CGGAATTGAC	900
	CTACCAGATG	AATCTACTGG	ATTTGTTCCC	AAAGAGTATA	GCTTTGCTAA	950
40	TTTCATTACC	AATGCCTTTG	GGCAGTTTGA	TAACTATACG	CCGATGCAGT	1000
	TGGCTCAGTA	TGTAGCAACT	ATTGCAAATG	ATGGTGTTCG	TGTGGCTCCT	1050
	CGTATTGTTG	AAGGCATTTA	TGGTAATAAT	GATAAGGGAG	GACTGGGTGA	1100
	CTTGATTGAG	CAACTGCAAC	CGACAGAGAT	GAATAAGGTC	AATATATCCG	1150
	ACTCCGATAT	GAGCATCTTG	CACCAAGGTT	TTTATCAGGT	TGCCCATGGT	1200
45	ACTAGTGAAT	TGACAACTGG	ACGTGCCTTT	TCAAATGGCG	CCTTGGTATC	1250
	CATTAGCGGA	AAAACAGGTA	CAGCCGAAAG	CTATGTGGCA	GATGGTCAGC	1300
	AAGCAACCAA	TACCAATGCG	GTGGCCTATG	CCCCATCTGA	TAATCCCCAA	1350
	ATCGCTGTAG	CTGTTGTCTT	CCCTCATAAC	ACCAACCTTA	CAAATGGTGT	1400
	CGGACCTTCC	ATTGCGCGCG	ATATTATCAA	CCTCTATAAC	CAACATCATC	1450
50	CAATGAATTA	GAAAGGAACA	TATGCTTTAT	CCAACACCTA	TTGCTAAGCT	1500
	AATTGACAGT	TATTCGAAGT	TACCGGGTAT	CGGGATTAAG	ACGGCTACCC	1550
	GTTTGGCCTT					1560

## 2) INFORMATION FOR SEQ ID NO: 1173

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2007 bases  
 (B) TYPE: Nucleic acid

(C) STRANDEDNESS: Double  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Streptococcus pneumoniae*  
(B) STRAIN: CS109  
(C) ACCESSION NUMBER: Z49096

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1173

	GAGTCCCGAT	TGCTGAGGAT	GCAACCTCCT	ATAATGTCTA	TGCGGTCATT	50
	GATGAGAACT	ATAAGTCAGC	AACGGGTAAG	ATTCTTTACG	TAGAAAAAAC	100
15	ACAATTTAAC	AAGGTTGCAG	AGGTCTTTCA	TAAGTATCTG	GACATGGAAG	150
	AATCCTATGT	AAGAGAGCAA	CTCTCGCAAC	CTAATCTCAA	GCAAGTTTCC	200
	TTTGGAGCAA	AGGGAAATGG	GATTACCTAT	GCCAATATGA	TGACTATCAA	250
	AAAAGAGTTG	GAAACTGCAG	AGGTCAAGGG	GATTGATTTT	ACAACCAGTC	300
	CCAATCGTAG	TTATCCAAAC	GGACAATTTG	CTTCTAGTTT	TATCGGTCTA	350
20	GCTCAGCTCC	ATGAAAATGA	AGATGGCAGC	AAGAGCTTGC	TGGGAACTTC	400
	TGGAATGGAG	AGTTCCTTGA	ACAGTATTCT	TGCAGGGACA	GACGGCATTG	450
	TTACCTATGA	AAAGGATCGT	CTGGGCAATA	TTGTACCCGG	AACAGAACTG	500
	GTATCGCAAC	AAACTGTGGA	TGGCAAGGAT	GTTTATACAA	CATTGTCTAG	550
	TCCGCTACAA	TCTTTTCATGG	AAACTCAGAT	GGATGCCTTT	CTAGAAAAAG	600
25	TAAAAGGTAA	GTATATGACC	GCGACCTTGG	TCAGTGCAAA	GACCGGTGAA	650
	ATTCTCGCTA	CCACCCAACG	ACCTACCTTT	AATGCAGATA	CTAAAGAAGG	700
	AATCACTGAG	GACTTTGTTT	GGCGTGATAT	TCTTTATCAA	AGTAACTATG	750
	AACCAGGATC	AGCCTTTAAG	GTCATGATGT	TAGCTTCTTC	TATTGATAAT	800
	AATACCTTCC	CAAGTGGAGA	ATACTTCAAT	AGCAGTGAAT	TCAAAAATAGC	850
30	GGATGCGACG	ACTCGAGATT	GGGATGTTAA	TGAGGGTTTG	ACTACTGGTG	900
	GGATGATGAC	TTTCTCACAA	GGTTTCGCTC	ACTCCAGTAA	TGTTGGAACG	950
	AGTCTACTTG	AACAAAAAAT	GGGAGATGCT	ACTTGGTTGG	ATTATCTAAA	1000
	ACGCTTTAAA	TTTGGGGTTC	CAACTCGCTT	TGGCTTGACA	GATGAATACG	1050
	CTGGTCAACT	TCCAGCTGAT	AATATTGTTA	GTATTGCTCA	AAGCTCATTT	1100
35	GGGCAAGGAA	TTTCAGTGAC	ACAAACACAA	ATGCTTCGTG	CCTTTACAGC	1150
	TATTGCTAAT	GATGGAGTTA	TGCTGGAGCC	AAAATTTATA	AGTGCTATTT	1200
	ATGATACTAA	CAATCAGTCT	GTACGTAAGT	CACAAAAAGA	AATAGTAGGA	1250
	AATCCTGTTT	CCAAAGAGGC	AGCAAGCACA	ACTCGAAATC	ACATGATCTT	1300
	AGTTGGGACG	GACCCCTCTAT	ATGGAACATAT	GTATAATCAC	TACACAGGAA	1350
40	AGCCAATTAT	AACAGTTCCT	GGACAAAATG	TAGCAGTTAA	ATCCCGTACG	1400
	GCTCAAAATCG	CTGATGAGAA	AAATGGAGGA	TACTTGGTTG	GTTCTACCAA	1450
	TTATATTTTC	TCAGTTGTGA	CTATGAATCC	TGCTGAAAAT	CCTGATTTTA	1500
	TCTTGATATG	AACGGTTCAA	CAGCCTGAGC	ATTATTCAGG	TATCCAGTTG	1550
	GGAGAATTTG	CCACCCCAAT	CTTGGAGCGG	GCTTCAGCTA	TGAAAGAATC	1600
45	TCTCAATCTT	CAATCTCCAG	CCAAAAATTT	AGATAAAGTT	ACGACAGAAT	1650
	CTTCTTATGC	AATGCCTAGC	ATCAAGGATA	TTTCACCTGG	TGAGTTGGCG	1700
	GAAGCCTTAC	GCCGAAATAT	TGTGCAACCA	ATCGTTGTAG	GTAAGTGAAC	1750
	AAAGATTAAA	GAGACTTCTG	TAGAAGAAGG	GACCAATCTT	GCACCAAACC	1800
	AACAAGTTCT	CCTTTTATCG	GATAAGGTAG	AAGAAATTCC	AGACATGTAT	1850
50	AGCTGGAAAA	AAGAGACTGC	CGAGACCTTT	GCTAAATGGT	TGGATATTGA	1900
	ACTGGAATTT	GAAGGTTTCAG	GTTCCGTTGT	TCAGAAGCAA	GATGTTTCGGA	1950
	CTAATACAGC	TATCAAAAAC	ATTAAAAAAA	TTAAATTAAC	TTTAGGAGAC	2000
	TAATATG					2007

2) INFORMATION FOR SEQ ID NO: 1174

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases

(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

5 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1174

GAACGTGGTG AAGTTCGC

18

10

2) INFORMATION FOR SEQ ID NO: 1175

15 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1175

25 GTTACTGGTG TAGAAATGTT C

21

2) INFORMATION FOR SEQ ID NO: 1176

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1176

40

TACTGGTGTA GAAATGTTC

19

2) INFORMATION FOR SEQ ID NO: 1177

45

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1177

55

GCTCAACAAG TTCCAGATTA

20

60

## 2) INFORMATION FOR SEQ ID NO: 1178

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2456 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus aureus*  
 (B) STRAIN: NCTC8325  
 (C) ACCESSION NUMBER: X52593

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1178

	ATGAACTGAT	TATACTTAAC	ATTAAAAAAG	ATGATAACAC	CTTCTACACC	50
	TCCATATCAC	AAAAAATTAT	AACATTATTT	TGACATAAAT	ACTACATTTG	100
20	TAATATACTA	CAAATGTAGT	CTTATATAAG	GAGGATATTG	ATGAAAAAGA	150
	TAAAAATTGT	TCCACTTATT	TTAATAGTTG	TAGTTGTCGG	GTTTGGTATA	200
	TATTTTATATG	CTTCAAAAGA	TAAAGAAATT	AATAATACTA	TTGATGCAAT	250
	TGAAGATAAA	AATTTCAAAC	AAGTTTATAA	AGATAGCAGT	TATATTTCTA	300
	AAAGCGATAA	TGGTGAAGTA	GAAATGACTG	AACGTCCGAT	AAAAATATAT	350
25	AATAGTTTAG	GCGTTAAAGA	TATAAACATT	CAGGATCGTA	AAATAAAAAA	400
	AGTATCTAAA	AAATAAAAAAC	GAGTAGATGC	TCAATATAAA	ATTAAAACAA	450
	ACTACGGTAA	CATTGATCGC	AACGTTCAAT	TTAATTTTGT	TAAAGAAGAT	500
	GGTATGTGGA	AGTTAGATTG	GGATCATAGC	GTCATTATTC	CAGGAATGCA	550
	GAAAGACCAA	AGCATAcata	TTGAAAATTT	AAAATCAGAA	CGTGGTAAAA	600
30	TTTTAGACCG	AAACAATGTG	GAATTGGCCA	ATACAGGAAC	ACATATGAGA	650
	TTAGGCATCG	TTCCAAGAG	TGTATCTAAA	AAAGATTATA	AAGCAATCGC	700
	TAAAGAACTA	AGTATTTCTG	AAGACTATAT	CAACAACAAA	TGGATCAAAA	750
	TTGGGTACAA	GATGATACCT	TCGTTCCACT	TTAAAACCGT	TAAAAAAATG	800
	GATGAATATT	TAAGTGATTT	CGCAAAAAAA	TTTCATCTTA	CAACTAATGA	850
35	AACAGAAAGT	CGTAACTATC	CTCTAGAAAA	AGCGACTTCA	CATCTATTAG	900
	GTTATGTTGG	TCCCATTAAC	TCTGAAGAAT	TAAAACAAAA	AGAATATAAA	950
	GGCTATAAAG	ATGATGTCAGT	TATTGGTAAA	AAGGGACTCG	AAAAACTTTA	1000
	CGATAAAAAAG	CTCCAACATG	AAGATGGCTA	TCGTGTCACA	ATCGTTGACG	1050
	ATAATAGCAA	TACAATCGCA	CATACATTAA	TAGAGAAAAA	GAAAAAAGAT	1100
40	GGCAAAGATA	TTCAACTAAC	TATTGATGCT	AAAGTTCAAA	AGAGTATTTA	1150
	TAACAACATG	AAAAATGATT	ATGGCTCAGG	TACTGCTATC	CACCC'TCAAA	1200
	CAGGTGAATT	ATTAGCACTT	GTAAGCACAC	CTTCATATGA	CGTCTATCCA	1250
	TTTATGTATG	GCATGAGTAA	CGAAGAATAT	AATAAATTAA	CCGAAGATAA	1300
	AAAAGAACCT	CTGCTCAACA	AGTTCCAGAT	TACAAC'TTCA	CCAGGTTCAA	1350
45	CTCAAAAAAT	ATTAACAGCA	ATGATTGGGT	TAAATAACAA	AACATTAGAC	1400
	GATAAAACAA	GTTATAAAAT	CGATGGTAAA	GGTTGGCAAA	AAGATAAATC	1450
	TTGGGGTGGT	TACAACGTTA	CAAGATATGA	AGTGGTAAAT	GGTAATATCG	1500
	ACTTAAACAA	AGCAATAGAA	TCATCAGATA	ACATTTTCTT	TGCTAGAGTA	1550
	GCACTCGAAT	TAGGCAGTAA	GAAATTTGAA	AAAGGCATGA	AAAAACTAGG	1600
50	TGTTGGTGAA	GATATACCAA	GTGATTATCC	ATTTTATAAT	GCTCAAATTT	1650
	CAAACAAAAA	TTTAGATAAT	GAAATATTAT	TAGCTGATTG	AGGTTACGGA	1700
	CAAGGTGAAA	TACTGATTAA	CCCAGTACAG	ATCCTTTTCAA	TCTATAGCGC	1750
	ATTAGAAAAAT	AATGGCAATA	TTAACGCACC	TCAC'TTATTA	AAAGACACGA	1800
	AAAACAAAGT	TTGGAAGAAA	AATATTATTT	CCAAAGAAAA	TATCAATCTA	1850
55	TTAAATGATG	GTATGCAACA	AGTCGTAAAT	AAAACACATA	AAGAAGATAT	1900
	TTATAGATCT	TATGCAAACT	TAATTTGGCAA	ATCCGGTACT	GCAGAACTCA	1950
	AAATGAAACA	AGGAGAAAAGT	GGCAGACAAA	TTGGGTGGTT	TATATCATAT	2000
	GATAAAGATA	ATCCAAACAT	GATGATGGCT	ATTAATGTTA	AAGATGTACA	2050
	AGATAAAGGA	ATGGCTAGCT	ACAATGCCAA	AATCTCAGGT	AAAGTGTATG	2100
60	ATGAGCTATA	TGAGAACGGT	AATAAAAAAT	ACGATATAGA	TGAATAACAA	2150

	AACAGTGAAG	CAATCCGTAA	CGATGGTTGC	TTCACTGTTT	TATTATGAAT	2200
	TATTAATAAG	TGCTGTTACT	TCTCCCTTAA	ATACAATTTC	TTCATTTTCA	2250
	TTGTATGTTG	AAAGTGACAC	TGTAACGAGT	CCATTTTCTT	TTTTTATGGA	2300
	TTTCTTATTT	GTAATTTTCAG	CGATAACGTA	CAATGTATTA	CCTGGTATAC	2350
5	AGTTTAATAA	ATTTAACGTT	ATTCATTTGT	GTTCTTGCTA	CAACTTCTTC	2400
	TCCGTATTTA	CCTTCTTCTA	CCCATAATTT	AAATGATATT	GAAAGTGTAT	2450
	GCATGC					2456

10

## 2) INFORMATION FOR SEQ ID NO: 1179

## (i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## 20 (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1179

ATTTGGTGAC GGGTGACTTT

20

25

## 2) INFORMATION FOR SEQ ID NO: 1180

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 19 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

## 35 (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1180

TCCACCGTTG CCAATCGCA

19

40

## 2) INFORMATION FOR SEQ ID NO: 1181

## (i) SEQUENCE CHARACTERISTICS:

- 45 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

50

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1181

55 AGCAGCTTAC TAGATGCCGT

20

60

## 2) INFORMATION FOR SEQ ID NO: 1182



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1182

AACTGCAAGA GATCCTTTGG

20

## 2) INFORMATION FOR SEQ ID NO: 1183

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2535 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: 175  
 (C) ACCESSION NUMBER: M18729

## (xii) SEQUENCE DESCRIPTION: SEQ ID NO: 1183

ATGGCGATAG	AAAAGCTATC	ACCCGGCATG	CAACAGTATG	TGGATATTAA	50
AAAGCAATAT	CCAGATGCTT	TTTTGCTCTT	TCGGATGGGT	GATTTTTATG	100
AATTATTTTA	TGAGGATGCG	GTCAATGCTG	CGCAGATTCT	GGAAATTTCC	150
TTAACGAGTC	GCAACAAGAA	TGCCGACAAT	CCGATCCCTA	TGGCGGGTGT	200
TCCCTATCAT	TCTGCCCAAC	AGTATATCGA	TGTCTTGATT	GAGCAGGGTT	250
ATAAGGTGGC	TATCGCAGAG	CAGATGGAAG	ATCCTAAACA	AGCAGTTGGG	300
GTTGTTAAAC	GAGAGGTTGT	TCAGGTCATT	ACGCCAGGGA	CAGTGGTCGA	350
TAGCAGTAAG	CCGGACAGTC	AGAATAATTT	TTTGGTTTCC	ATAGACCGCG	400
AAGGCAATCA	ATTTGGCCTA	GCTTATATGG	ATTTGGTGAC	GGGTGACTTT	450
TATGTGACAG	GTCTTTTGGA	TTTCACGCTG	GTTTGTGGGG	AAATCCGTAA	500
CCTCAAGGCT	CGAGAAGTGG	TGTTGGGTTA	TGACTTGTCT	GAGGAAGAAG	550
AACAAATCCT	CAGCCGCCAG	ATGAATCTGG	TACTCTCTTA	TGAAAAAGAA	600
AGCTTTGAAG	ACCTTCATTT	ATTGGATTG	CGATTGGCAA	CGGTGGAGCA	650
AACGGCATCT	AGTAAGCTGC	TCCAGTATGT	TCATCGGACT	CAGATGAGGG	700
AATTGAACCA	CCTCAAACCT	GTTATCCGCT	ACGAAATTAA	GGATTTCTTG	750
CAGATGGATT	ATGCGACCAA	GGCTAGTCTG	GATTTGGTTG	AGAATGCTCG	800
CTCAGGTAAG	AAACAAGGCA	GTCTTTTCTG	GCTTTTGGAT	GAAACCAAAA	850
CGGCTATGGG	GATGCGTCTC	TTGCGTTCCT	GGATTCATCG	CCCCTTGATT	900
GATAAGGAAC	GAATCGTCCA	ACGTCAAGAA	GTAGTGCAGG	TCTTTCTCGA	950
CCATTTCTTT	GAGCGTAGTG	ACTTGACAGA	CAGTCTCAAG	GGTGTTTATG	1000
ACATTGAGCG	CTTGGCTAGT	CGTGTTTCTT	TTGGCAAAAC	CAATCCAAAG	1050
GATCTCTTGC	AGTTGGCGAC	TACCTTGTCT	AGTGTGCCAC	GGATTCGTGC	1100
GATTTTAGAA	GGGATGGAGC	AACCTACTCT	AGCCTATCTC	ATCGCACAAAC	1150
TGGATGCAAT	CCCTGAGTTG	GAGAGTTTGA	TTAGCGCAGC	GATTGCTCCT	1200
GAAGCTCCTC	ATGTGATTAC	AGATGGGGGA	ATTATCCGGA	CTGGATTTGA	1250
TGAGACTTTA	GACAAGTATC	GTTGCGTTCT	CAGAGAAGGG	ACTAGCTGGA	1300
TTGCTGAGAT	TGAGGCTAAG	GAGCGAGAAA	ACTCTGGTAT	CAGCACGCTC	1350
AAGATTGACT	ACAATAAAAA	GGATGGCTAC	TATTTTCATG	TGACCAATTC	1400

## DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET  
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 3 DE 4

NOTE: Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

## JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE  
THAN ONE VOLUME

THIS IS VOLUME 3 OF 4

NOTE: For additional volumes please contact the Canadian Patent Office

## DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET  
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 4 DE 4

NOTE: Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

---

## JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE  
THAN ONE VOLUME

THIS IS VOLUME 4 OF 4

NOTE: For additional volumes please contact the Canadian Patent Office

```

GCAACTGGGA AATGTGCCAG CCCACTTTTT CCGCAAGGCG ACGCTGAAAA 1450
ACTCAGAACG CTTTGGAACC GAAGAATTAG CCCGTATCGA GGGAGATATG 1500
CTTGAGGCGC GTGAGAAGTC AGCCAACCTC GAATACGAAA TATTTATGCG 1550
CATTCGTGAA GAGGTCGGCA AGTACATCCA GCGTTTACAA GCTCTAGCCC 1600
5 AAGGAATTGC GACGGTTGAT GTCTTACAGA GTCTGGCGGT TGTGGCTGAA 1650
ACCCAGCATT TGATTCGACC TGAGTTTGGT GACGATTCAC AAATTGATAT 1700
CCGGAAAGGG CGCCATGCTG TCGTTGAAAA GGTATGGGG GCTCAGACCT 1750
ATATTCCAA TACGATTCAG ATGGCAGAAG ATACCAGTAT TCAATTGGTT 1800
ACAGGGCCAA ACATGAGTGG GAAGTCTACC TATATGCGTC AGTTAGCCAT 1850
10 GACGGCGGTT ATGGCCCAGC TGGGTTCCTA TGTTCTTGCT GAAAGCGCCC 1900
ATTTACCGAT TTTTGATGCG ATTTTACCC GTATCGGAGC AGCAGATGAC 1950
TTGGTTTCGG GTCAGTCAAC CTTTATGGTG GAGATGATGG AGGCCAATAA 2000
TGCCATTTTCG CATGCGACCA AGAAGTCTCT CATTCTCTTT GATGAATTGG 2050
GACGTGGAAC TGCAACTTAT GACGGGATGG CTCTTGCTCA GTCCATCATC 2100
15 GAATATATCC ATGAGCACAT CGGAGCTAAG ACCCTCTTTG CGACCCACTA 2150
CCATGAGTTG ACTAGTCTGG AGTCTAGTTT ACAACACTTG GTCAATGTCC 2200
ACGTGGCAAC TTTGGAGCAG GATGGGCAGG TCACCTTCCT TCACAAGATT 2250
GAACCGGGAC CAGCTGATAA ATCCTACGGT ATCCATGTTG CCAAGATTGC 2300
TGGCTTGCCA GCAGACCTTT TAGCAAGGGC GGATAAGATT TTGACTCAGC 2350
20 TAGAGAATCA AGGAACAGAG AGTCCCTCCT CCATGAGACA AACTAGTGCT 2400
GTCACTGAAC AGATTTCAT CTTTGATAGG GCAGAAGAGC ATCCTATCCT 2450
AGCAGAATTA GCTAAACTGG ATGTGATATA TATGACACCT ATGCAGGTTA 2500
TGAATGTCTT AGTAGAGTTA AAACAGAAAC TATAA 2535

```

25

## 2) INFORMATION FOR SEQ ID NO: 1184

## (i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 623 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

35

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-05

## 40 (xiii) SEQUENCE DESCRIPTION: SEQ ID NO: 1184

```

TGACGGGTGA CTTTTATGTG ACAGGTCTTT TGGATTTCAC GCTGGTTTGT 50
GGGGAATCC GTAACCTCAA GGCTCGAGAA TGGGTGTGG GTTATGACTT 100
GTCTGAGGAA GAAGAACAAA TCCTCAGCCG CCAGATGAAT CTGGTACTCT 150
45 CTTATGAAAA AGAAAGCTTT GAAGACCTTC ATTTATTGGA TTTGCGATTG 200
GCAACGGTGG AGCAAACGGC ATCTAGTAAG CTGCTCCAGT ATGTTTCATCG 250
GACTCAGATG AGGGAATTGA ACCACCTCAA ACCTGTTATC CGATACGAAA 300
TTAAGGATTT CTTGCAGATG GATTATGCGA CCAAGGCTAG TCTGGATTTG 350
GTTGAGAATG CTCGCTCAGG TAAGAAACAA GGCAGTCTTT TCTGGCTTTT 400
50 GGATGAAACC AAAACGGCTA TGGGGATGCG TCTCTTGCGT TCTTGGATTC 450
ATCGCCCTTT GATTGATAAG GAACGAATCG TCCAACGTCA AGAAGTAGTG 500
CAGGTCTTTC TCGACCATTT CTTTGAGCGT AGTGACTTGA CAGACAGTCT 550
CAAGGGTGTT TATGACATTG AGCGCTTGGC TAGTCGTGTT TCTTTTGGCA 600
AAACCAATCC AAAGGATCTC TTG 623

```

55

## 2) INFORMATION FOR SEQ ID NO: 1185

## 60 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 621 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-06

(xiv) SEQUENCE DESCRIPTION: SEQ ID NO: 1185

15	TGACGGGTGA	CTTTTATGTG	ACAGGTCTTT	TGGATTTCAC	GCTGGTTTGT	50
	GGGGAAATCC	GTAACCTCAA	GGCTCGAGAA	GTGGTGTGG	GTTATGACTT	100
	GTCTGAGGAA	GAAGAACAAA	TCCTCAGCCG	CCAGATGAAT	CTGGTACTCT	150
	CTTATGAAAA	AGAAAGCTTT	GAAGACCTTC	ATTTATTGGA	TTTGCGATTG	200
	GCAACGGTGG	AGCAAACGGC	ATCTAGTAAG	CTGCTCCAGT	ATGTTTCATCG	250
	GACTCAGATG	AGGGAATTGA	ACCACCTCAA	ACCTGTTATC	CGATACGAAA	300
20	TTAAGGATTT	CTTGCAGATG	GATTATGCGA	CCAAGGCTAG	TCTGGATTG	350
	GTTGAGAATG	CTCGCTCAGG	TAAGAAACAA	GGCAGTCTTT	TCTGGCTTTT	400
	GGATGAAACC	AAAACGGCTA	TGGGGATGCG	TCTCTTGCGT	TCTTGATTTC	450
	ATCGCCCCCTT	GATTGATAAG	GAACGAATCG	TCCAACGTCA	AGAAGTAGTG	500
	CAGGTCTTTC	TCGACCATTT	CTTTGAGCGT	AGTGACTTGA	CAGACAGTCT	550
25	CAAGGGTGTT	TATGACATTG	AGCGCTTGGC	TAGTCGTGTT	TCTTTTGGCA	600
	AAACCAATCC	AAAGGATCTC	T			621

30 2) INFORMATION FOR SEQ ID NO: 1186

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 622 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-11

(xv) SEQUENCE DESCRIPTION: SEQ ID NO: 1186

45	TGACGGGTGA	CTTTTATGTG	ACAGGTCTTT	TGGATTTCAC	GCTGGTTTGT	50
	GGGGAAATCC	GTAACCTCAA	GGCTCGAGAA	GTGGTGTGG	GTTATGACTT	100
	GTCTGAGGAA	GAAGAACAAA	TCCTCAGCCG	CCAGATGAAT	CTGGTACTCT	150
	CTTATGAAAA	AGAAAGCTTT	GAAGACCTTC	ATTTATTGGA	TTTGCGATTG	200
	GCAACGGTGG	AGCAAACGGC	ATCTAGTAAG	CTGCTCCAGT	ATGTTTCATCG	250
50	GACTCAGATG	AGGGAATTGA	ACCACCTCAA	ACCTGTTATC	CGCTACGAAA	300
	TTAAGGATTT	CTTGCAGATG	GATTATGCGA	CCAAGGCTAG	TCTGGATTG	350
	GTTGAGAATG	CTCGCTCAGG	TAAGAAACAA	GGCAGTCTTT	TCTGGCTTTT	400
	GGATGAAACC	AAAACGGCTA	TGGGGATGCG	TCTCTTGCGT	TCTTGATTTC	450
55	ATCGCCCCCTT	GATTGATAAG	GAACGAATCG	TCCAACGTCA	AGAAGTAGTG	500
	CAGGTCTTTC	TCGACCATTT	CTTTGAGCGT	AGTGACTTGA	CAGACAGTCT	550
	CAAGGGTGTT	TATGACATTG	AGCGCTTGGC	TAGTCGTGTT	TCTTTTGGCA	600
	AAACCAATCC	AAAGGATCTC	TT			622

## 2) INFORMATION FOR SEQ ID NO: 1187

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 622 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pneumoniae*  
 (B) STRAIN: StrR-55

## (ix) SEQUENCE DESCRIPTION: SEQ ID NO: 1187

	TGACGGGTGA	CTTTTATGTG	ACAGGTCTTT	TGGATTTCAC	GCTGGTTTGT	50
	GGGGAAATCC	GTAACCTCAA	GGCTCGAGAA	GTGGTGTTGG	GTTATGACTT	100
20	GTCTGAGGAA	GAAGAACAAA	TCCTCAGCCG	CCAGATGAAT	CTGGTACTCT	150
	CTTATGAAAA	AGAAAGCTTT	GAAGACCTTC	ATTTATTGGA	TTTGCGATTG	200
	GCAACGGTGG	AGCAAACGGC	ATCTAGTAAG	CTGCTCCGGT	ATGTTTCATCG	250
	GACTCAGATG	AGGGAATTGA	ACCACCTCAA	ACCTGTTATC	CGCTACGAAA	300
	TTAAGGATTT	CTTGCAGATG	GATTATGCGA	CCAAGGCTAG	TCTGGATTTG	350
25	GTTGAGAATG	CTCGCTCAGG	TAAGAAACAA	GGCAGTCTTT	TCTGGCTTTT	400
	GGATGAAACC	AAAACGGCTA	TGGGGATGCG	TCTCTTGCGT	TCTTGGATTC	450
	ATCGCCCCTT	GATTGATAAG	GAACGAATCG	TCCAACGTCA	AGAAGTAGTG	500
	CAGGTCTTTC	TCGACCATTT	CTTTGAGCGT	AGTGACTTGA	CAGACAGTCT	550
	CAAGGGTGTT	TATGACATTG	AGCGCTTGGC	TAGTCGTGTT	TCTTTTGGCA	600
30	AAACCAATCC	AAAGGATCTC	TT			622

## 2) INFORMATION FOR SEQ ID NO: 1188

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 599 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus oralis*  
 (B) STRAIN: ATCC 35037

## (ix) SEQUENCE DESCRIPTION: SEQ ID NO: 1188

50	GGGTGACTTT	TATGTAACGG	GGCTATTGGA	TTTCACGTTG	GTTTGTGGGG	50
	AAATTCGCAA	TCTCAAGGCT	AGAGAAGTGG	TGCTGGGTTA	TGACTTGTCT	100
	GAGGAAGAAG	AACAAATCCT	CAGTCGTCAG	ATGAATCTGG	TGCTTTCTTA	150
	TGAGAAGGAA	GGCTTTGAGG	ACCTTCATTT	ACTGGATCCA	CGACTGGCAG	200
	CTGTGGAGCA	AGCGGCAGCT	AGTAAGCTCC	TCCAGTATGT	TCACCGGACC	250
55	CAGATGCGGG	AATTGAACCA	CCTCAAACCA	GTTATCCGCT	ATGAAATCAA	300
	AGATTTCTTA	CAGATGGACT	ATGCGACCAA	GGCTAGTCTG	GATTTGGTTG	350
	AGAAATGCCC	GTCAGGCAAG	AAGCAAGGCA	GTCTTTTCTG	GCTTTTAGAT	400
	GAAACCAAGA	CGGCTATGGG	AATGCGTCTC	TTGCGTTCTT	GGATTTCATCG	450
	TCCTTTGATT	GATAAGGAGC	GAATCGTCCA	ACGTCAAGAG	GTGGTGCAGG	500
60	TCTTTCTTGA	CCACTTCTTT	GAGCGTAGTG	ATTTAACGGA	CAGTCTTAAG	550

GGTGTATTATG ATATCGAACG CTTGGCTAGT CGGGTTTCTT TTGGCAAGA

599

## 5 2) INFORMATION FOR SEQ ID NO: 1189

## (i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 624 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## 15 (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
 (B) STRAIN: ATCC 49456

## (x) SEQUENCE DESCRIPTION: SEQ ID NO: 1189

20 GGTGACGGGT GACTTTTATG TGACAGGTCT TTTGGATTTT ACGCTGGTTT 50  
 GTGGGGAAAT CCGCAATCTC AAGGCTCGAG AAGTGGTGCT GGGTTATGAC 100  
 TTGTCTGAGG AAGAAGAACA GATCCTTAGT CGTCAGATGA ATCTGGTACT 150  
 TTCCTATGAA AAAGAAGGCT TTGAAGACCT TCATTTACTG GATTCACGAT 200  
 25 TGGCAGCTGT GGAGCAAGCG GCATCTAGTA AACTGCTTCA GTATGTTTCA 250  
 CGGACTCAGA TGAGGGAATT GAACCACCTC AAGCCTGTTA TCCGCTATGA 300  
 AATCAAAGAT TTTTTCGAGA TGGATTATGC GACCAAGGCT AGTCTGGATT 350  
 TGGTTGAGAA TGCCCGTTCA GGCAAGAAGC AAGGTAGTCT TTTTGTGGCTT 400  
 TTGGATGAAA CCAAACAGC TATGGGAATG CGTCTCTTGC GGTCTTGGAT 450  
 30 TCATCGCCCC CTGATTGATA AGGAACGAAT TGTCCAACGC CAAGAAGTTG 500  
 TGCAGGTCTT TCTCGACCAT TTCTTTGAGC GTAGTGATTT GACAGACAGT 550  
 CTCAAGGGTG TTTATGACAT TGAGCGCTTG GCTAGTCGTG TTTCTTTTGG 600  
 CAAAACCAAT CCAAAGGATC TCTT 624

35

## 2) INFORMATION FOR SEQ ID NO: 1190

## (i) SEQUENCE CHARACTERISTICS:

- 40 (A) LENGTH: 599 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## 45 (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus mitis*  
 (B) STRAIN: LSPQ 2583

50

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1190

55 TGACGGGTGA CTTTCAGGTG ACTAGTTTAG AGGACTTTGT CTTGGTCTGC 50  
 GGGGAAATCC GCAATTGAA AGCTAGGGAA GTGGTGCTGG GCTATGCCTT 100  
 GCCAGAAGCT GAGGAGCAGG TTTTGGCTGG ACAGATGAAC CTTTACTGT 150  
 CCTATGTGGA GAAGGTTTGT GAGGATGTTT AGCTGCTGGG CGAGGAGCTG 200  
 TCTCCTATGG AGCGTCAGGC AGCAGGGAAA CTGCTGGAGT ATGTGCACCG 250  
 GACCCAGATG AGGGAGCTCA GCCATTTGAA GAAGGCTCAG CATTATGAAA 300  
 TCAAGGACTT CCTGCAAATG GACTATGCCA CCAAGGCGAG TCTGGATTTG 350  
 60 ACAGAAAATG CTCGCTCGGG CAAGAAGCAC GGCAGTCTTT ATTGGCTGAT 400

GGACGAGACT	AAGACGGCCA	TGGGCGGCCG	CATGCTGCGC	TCTTGGATCC	450
AGCGTCCGCT	GATTGATGAA	GCGCGAATTA	GCCAGCGACA	GAATGTCGTT	500
GAGGTTTTTC	TGGATCATTT	CTTTGAGCGG	AGTGATTTGA	CGGAGAGCCT	550
CAAGGGGGTC	TATGATATCG	AGCGGCTGGC	TAGTCGGGTG	TCTTTTGGC	599

5

## 2) INFORMATION FOR SEQ ID NO: 1191

- 10 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 622 bases
  - (B) TYPE: Nucleic acid
  - (C) STRANDEDNESS: Double
  - (D) TOPOLOGY: Linear
- 15 (ii) MOLECULE TYPE: Genomic DNA
- (vi) ORIGINAL SOURCE:
- (A) ORGANISM: *Streptococcus mitis*
  - (B) STRAIN: ATCC 903
- 20 (ix) SEQUENCE DESCRIPTION: SEQ ID NO: 1191

25	TGACGGGTGA	CTTTCAGGTG	ACTAGTTTAG	AGGACTTTGC	CCTGGTCTGC	50
	GGGGAAATCC	GTAATTTGAA	GGCTAGGGAA	GTGGTGCTGG	GCTATGCTTT	100
	GCCAGAAGCT	GAGGAGCAGG	TCTTGGCTGG	ACAGATGAAT	CTTTTGCTGT	150
	CCTATGTACA	GACGGCCTTG	GACGATGTCC	AGCTGCTGGG	CGAGGAACTG	200
	TCTCCTATGG	AGCGTCAGGC	AGCGGGGAAA	TTGCTAGAGT	ATGTGCACCG	250
	GACCCAGATG	AGGGAGCTCA	GCCATTTGAA	GAAGGCCAG	CATTATGAAA	300
30	TCAAGGACTT	TCTGCAAATG	GATTATGCTA	CCAAGGCGAG	TCTGGATTG	350
	ACAGAAAATG	CTCGCTCGGG	TAAGAAACAC	GGCAGTCTTT	ATTGGCTGAT	400
	GGACGAGACC	AAGACGGCCA	TGGGCGGCCG	TATGCTGCGC	TCTTGGATCC	450
	AGCGTCCGTT	GATTGATGAA	GTGCGAATTA	GCCAGCGGCA	GAATGTCGTC	500
	GAGGTTTTTC	TGGAACATTT	CTTTGAGCGG	AGTGATTTGA	CGGAGAGCCT	550
35	CAAGGGAGTC	TATGATATCG	AGCGGCTGGC	TAGTCGGGTG	TCTTTTGGCA	600
	AGACCAATCC	AAAGGATCTC	TT			622

## 40 2) INFORMATION FOR SEQ ID NO: 1192

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 bases
  - (B) TYPE: Nucleic acid
  - 45 (C) STRANDEDNESS: Single
  - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

## 50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1192

GGTAAAACAG GAACCTCTAA CT

22

55

## 2) INFORMATION FOR SEQ ID NO: 1193

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 bases
  - 60 (B) TYPE: Nucleic acid

855



(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1193

GGTAAGACAG GTACTTCTAA CT

22

10

2) INFORMATION FOR SEQ ID NO: 1194

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

20 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1194

25 CATTTC AAGT AATACAACAG AATC

24

30

2) INFORMATION FOR SEQ ID NO: 1195

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1195

40

CATTTC AAGT AACACA ACTG AATC

24

45 2) INFORMATION FOR SEQ ID NO: 1196

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1196

GCCATTTC AA GTAATACAAC AGAA

24

60

## 2) INFORMATION FOR SEQ ID NO: 1197

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1197

CAAACGCCAT TTCAAGTAAT ACAAC

25

## 2) INFORMATION FOR SEQ ID NO: 1198

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 381 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus saprophyticus*
- (B) STRAIN: ATCC 43867

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1198

AACGGGCGTC	TCGATAGAAA	AACACGTGAA	AATCCCAATG	ATTATAAACA	50
ATCAATATAC	GATTTTGCTG	AAGCTGTAAC	AAAAGGTATT	AAGGAACAAA	100
CAAATAAAAA	TTAATAGGCA	ACTTAACCAG	AATCGTTAAA	ACTATATGAC	150
GATTCTGGTT	TTTTAAATTC	AAAAAGTTTT	CTAAAAAATT	TACTTGCTTC	200
TTTAAAGTAT	AGGTATGAAA	TACAATTGAT	TAAAATAGTA	AAGGAAATGA	250
ATCATGAAAC	AATTAATAA	GCCTTTATAC	TTTACCTAT	TACTTTTAT	300
TACAACAACG	CTGATTGGCG	CGTTACTATT	ATATTTGCCA	ATCACAKGTA	350
AACATCCTAT	TGATTTTGTG	GACGCCCGTT	A		381

## 2) INFORMATION FOR SEQ ID NO: 1199

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1199

GTATTAAAGA AGATATCCAA AAAGC

25

## 2) INFORMATION FOR SEQ ID NO: 1200

857

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1200  
TCAAAGAAGA AACTAAAAA GCTGT

2) INFORMATION FOR SEQ ID NO: 1201

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1201  
AACGTAGGTG TCCTTCTTC

2) INFORMATION FOR SEQ ID NO: 1202

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1202  
GTGTTGAAAT GTTCCGTAAA CA

2) INFORMATION FOR SEQ ID NO: 1203

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 21 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1203  
GGIGARMGIG GIAAYGARAT G

## 2) INFORMATION FOR SEQ ID NO: 1204

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1204

GCIAAYAACI TCIWMYATGC C

21

## 2) INFORMATION FOR SEQ ID NO: 1205

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1205

AAYACITCIA WYATGCCIGT

20

## 2) INFORMATION FOR SEQ ID NO: 1206

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1206

CKISRIGTIG ARTCIGCCA

19

## 2) INFORMATION FOR SEQ ID NO: 1207

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

859

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1207

CCITCITCWC CIGGCATYTC

20

5

2) INFORMATION FOR SEQ ID NO: 1208

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1208

20 TCAAAAAGTT TTCTAAAAAA TTTAC

25

2) INFORMATION FOR SEQ ID NO: 1209

25

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 25 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1209

35

ACGGGCGTCC ACAAATCAA TAGGA

25

40 2) INFORMATION FOR SEQ ID NO: 1210

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 22 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

45

(ii) MOLECULE TYPE: DNA

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1210

ACCAGCTTGC CCAATACAAA GG

22

55

2) INFORMATION FOR SEQ ID NO: 1211

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 24 bases  
(B) TYPE: Nucleic acid

60

860

(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1211

ATTCTTGTA CAGGCTTTGA TCCC

24

2) INFORMATION FOR SEQ ID NO: 1212

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1212

CCICCCIRGIG GIGAIACIGC WCC

23

2) INFORMATION FOR SEQ ID NO: 1213

(i) SEQUENCE CHARACTERISTICS:

~~(A) LENGTH: 26 bases~~  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1213

AARGGIGGIA CIGCIGCIAT HCCIGG

26

2) INFORMATION FOR SEQ ID NO: 1214

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1214

GGTAAAACAG GTACCTCTAA CTA

23

2) INFORMATION FOR SEQ ID NO: 1215

861

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1337 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Streptococcus pyogenes*  
 (B) STRAIN: D471  
 (C) ACCESSION NUMBER: X65717

## (xii) SEQUENCE DESCRIPTION: SEQ ID NO: 1215

	AACAAAATAA	AAGAACTTAC	CTATTTTCCA	TCCAAAATGT	TTAGCAATCA	50
	TCATCTGCAA	GGCAACGTAT	TGCATGGCAT	TGATGTGATG	AGCAACTAAT	100
	ATGTCATTAG	AACGTTGCGT	CAAAC TAGCA	TCTAAATAAA	GATCGAAATG	150
20	CAGTTATCAA	AAATGCAAGC	TCCTATCGGC	CCTTGTTTTA	ATTATTACTC	200
	ACATTGCCTT	AATGTATTTA	CTTGCTTATT	ATTAAC TTTT	TTGCTAAGTT	250
	AGTAGCGTCA	GTTATTCATT	GAAAGGACAT	TATTATGAAA	ATTCTTGTA	300
	CAGGCTTTGA	TCCCTTTGGC	GGCGAAGCTA	TTAATCCTGC	CCTTGAAGCT	350
	ATCAAGAAAT	TGCCAGCAAC	CATTCATGGA	GCAGAAATCA	AATGTATTGA	400
25	AGTTCCAACG	GTTTTTCAAA	AATCTGCCGA	TGTGCTCCAG	CAGCATATCG	450
	AAAGCTTTCA	ACCTGATGCA	GTCCTTTGTA	TTGGGCAAGC	TGGTGGCCGG	500
	ACTGGACTAA	CGCCAGAACG	CGTTGCCATT	AATCAAGACG	ATGCTCGCAT	550
	TCCTGATAAC	GAAGGGAATC	AGCCTATTGA	TACACCTATT	CGTGCAGATG	600
	GTAAAGCAGC	TTATTTTTCA	ACCTTGCCAA	TCAAAGCGAT	GGTTGCTGCC	650
30	ATTCATCAGG	CTGGGCTTCC	TGCTTCTGTT	TCTAATACAG	CTGGTACCTT	700
	TGTTTGCAAT	CATTTGATGT	ATCAAGCCCT	TTACTTAGTG	GATAAATATT	750
	GTCCAATGTC	CAAAGCTGGG	TTTATGCATA	TTCCCTTTAT	GATGGAACAG	800
	GTGTTTGATA	AACCTAATAC	AGCTGCCATG	AACCTCGATG	ATATTACAAG	850
	AGGAATTGAG	GCTGCTATTT	TTGCCATTGT	CGATTTCAAA	GATCGTTCCG	900
35	ATTTAAAACG	TGTAGGGGGC	GCTACTCACT	GACTGTGACG	CTACTAAACC	950
	TATTTTAAAA	AAACAGAGAT	ATGAACTAAC	TCTGTTTTTT	TTGTGCTAAA	1000
	AATGAAAGAC	CTAGGGAAAC	TTTTTCATCGG	TCTTTCTCAA	TTGTCATCTT	1050
	AATCTAATA	TACTTCTAAC	ATCAGCGGGT	ATAGTTTGCC	AGTAATTAAG	1100
	AAACGTTGTT	GATCTAAATG	AGCAATCCCA	TTCAAAACAT	TAAGGTCAGG	1150
40	GTAAATGGGAC	TTATCAAGAT	TTAAGGCTTT	TAACAAAGGA	CTAATATCAT	1200
	AGGTGGCTAC	CACCTTTCCA	GAATCAGGTT	GGAGTTTGAC	AATAGTATTG	1250
	GTTTGCCAAA	TATTGGCATA	GAGATAACCA	TCTACATACT	CTAATTCGTT	1300
	AAGCATTGAG	ATAGGGACAC	TTTCTATAGC	AACTAGT		1337

## 2) INFORMATION FOR SEQ ID NO: 1216

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1216

GGTAAGACTG GTACATCAAA CTA

23

## 2) INFORMATION FOR SEQ ID NO: 1217

5 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1217  
15 CAAATGCCAT TTCAAGTAAC ACAAC 25

## 2) INFORMATION FOR SEQ ID NO: 1218

20 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
    (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1218  
30 CAAACGCCAT TTCAAGTAAC ACAAC 25

## 2) INFORMATION FOR SEQ ID NO: 1219

35 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
    (B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
    (ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1219  
CAAATGCTAT TTCAAGTAAT ACAAC 25

## 2) INFORMATION FOR SEQ ID NO: 1220

50 (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 25 bases  
55 (B) TYPE: Nucleic acid  
    (C) STRANDEDNESS: Single  
    (D) TOPOLOGY: Linear  
60 (ii) MOLECULE TYPE: DNA



(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1220

CAAACGC CATTTCAAGT AATACGAC

25

5

2) INFORMATION FOR SEQ ID NO: 1221

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

10

(ii) MOLECULE TYPE: DNA

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1221

GAYACICCG GICAYGTIGA YTT

23

20

2) INFORMATION FOR SEQ ID NO: 1222

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

25

(ii) MOLECULE TYPE: DNA

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1222

ATYGAYACIC CIGGICAYGT IGAYTT

26

35

2) INFORMATION FOR SEQ ID NO: 1223

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

40

(ii) MOLECULE TYPE: DNA

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1223

50

AYITCIARRT GIARYTCRCC CATICC

26

2) INFORMATION FOR SEQ ID NO: 1224

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

60

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1224  
CCIGYIHTIY TIGARCCIAT IATG

24

10

2) INFORMATION FOR SEQ ID NO: 1225

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1225

TAICCRAACA TYTCISMIAR IGGIAC

26

25

2) INFORMATION FOR SEQ ID NO: 1226

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear

35 (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1226

GTIRMRTAIC CRAACATYTC

40

20

2) INFORMATION FOR SEQ ID NO: 1227

45 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
50 (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1227

55 GTICCIYTIK CIGARATGTT YGGITA

26

60

2) INFORMATION FOR SEQ ID NO: 1228

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1228

GTICCIYTIK CIGARATGTT YGGITAYGC

29

## 2) INFORMATION FOR SEQ ID NO: 1229

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1229

TCCATYTGIG CIGCICCI GT IATCAT

26

## 2) INFORMATION FOR SEQ ID NO: 1230

## ----- (i) SEQUENCE CHARACTERISTICS: -----

- (A) LENGTH: 2145 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Escherichia coli*  
 (D) ACCESSION NUMBER: X00415

## (xiii) SEQUENCE DESCRIPTION: SEQ ID NO: 1230

45	TGAACGCCTA AAAGATAAAC GAGGAAACAA ATGGCTCGTA CAACACCCAT	50
	CGCACGCTAC CGTAACATCG GTATCAGTGC GCACATCGAC GCCGGTAAAA	100
	CCACTACTAC CGAACGTATT CTGTTCTACA CCGGTGTAAA CCATAAAATC	150
	GGTGAAGTTC ATGACGGCGC TGCAACCATG GACTGGATGG AGCAGGAGCA	200
	GGAACGTGGT ATTACCATCA CTTCCGCTGC GACTACTGCA TTCTGGTCTG	250
50	GTATGGCTAA GCAGTATGAG CCGCATCGCA TCAACATCAT CGACACCCCG	300
	GGGCACGTTG ACTTCACAAT CGAAGTAGAA CGTTCCATGC GTGTTCTCGA	350
	TGGTGCGGTA ATGGTTTACT GCGCAGTTGG TGGTGTTTCAG CCGCAGTCTG	400
	AAACCGTATG GCGTCAGGCA AACAAATATA AAGTTCCGCG CATTGCGTTC	450
	GTTAACAAAA TGGACCGCAT GGGTGCGAAC TTCTTGAAAG TTGTTAACCA	500
55	GATCAAAACC CGTCTGGGCG CGAACCCGGT TCCGCTGCAG CTGGCGATTG	550
	GTGCTGAAGA ACATTTACAC GTGTGTGTTG ACCTGGTGAA AATGAAAGCT	600
	ATCAACTGGA ACGACGCTGA CCAGGGCGTA ACCTTCGAAT ACGAAGATAT	650
	CCCGGCAGAC ATGGTTGAAC TGGCTAACGA ATGGCACCAG AACCTGATCG	700
	AATCCGCAGC TGAAGCTTCT GAAGAGCTGA TGGAAAAATA CCTGGGTGGT	750
60	GAAGAACTGA CTGAAGCAGA AATCAAAGGT GCTCTGCGTC AGCGCGTTCT	800

	GAACAACGAA	ATCATCCTGG	TAACCTGTGG	TTCTGCGTTC	AAGAACAAAG	850
	GTGTTTCAGGC	GATGCTGGAT	GCGGTAATTG	ATTACCTGCC	ATCCCCGGTT	900
	GACGTACCTG	CGATCAACGG	TATCCTGGAC	GACGGTAAAG	ACACTCCGGC	950
5	TGAACGTCAC	GCAAGTGATG	ACGAGCCGTT	CTCTGCACTG	GCGTTCAAAA	1000
	TCGCTACCGA	CCCGTTTGTT	GGTAACCTGA	CCTTCTTCCG	TGTTTACTCC	1050
	GGTGTGGTTA	ACTCTGGTGA	TACCGTACTG	AACTCCGTGA	AAGCTGCACG	1100
	TGAGCGTTTC	GGTCGTATCG	TTCAGATGCA	CGCTAACAAA	CGTGAAGAGA	1150
	TCAAAGAAGT	TCGCGCGGGC	GACATCGCTG	CTGCTATCGG	TCTGAAAGAC	1200
10	GTAACCACTG	GTGACACCCT	GTGTGACCCG	GATGCGCCGA	TCATTCTGGA	1250
	ACGTATGGAA	TTCCCTGAGC	CGGTAATCTC	CATCGCAGTT	GAACCGAAAA	1300
	CCAAAGCTGA	CCAGGAAAAA	ATGGGTCTGG	CTCTGGGCCG	TCTGGCTAAA	1350
	GAAGACCCGT	CTTTCCGTGT	ATGGACTGAC	GAAGAATCTA	ACCAGACCAT	1400
	CATCGCGGGT	ATGGGCGAAC	TGCACCTCGA	CATCATCGTT	GACCGTATGA	1450
	AGCGTGAATT	CAACGTTGAA	GCGAACGTAG	GTAAACCGCA	GGTTGCTTAC	1500
15	CGTGAAACTA	TCCGCCAGAA	AGTTACCGAT	GTTGAAGGTA	AACACGCGAA	1550
	ACAGTCTGGT	GGTCGTGGTC	AGTATGGTCA	TGTTGTTATC	GACATGTACC	1600
	CGCTGGAGCC	GGGTTCAAAC	CCGAAAGGCT	ACGAGTTCAT	CAACGACATT	1650
	AAAGGTGGTG	TAATCCCTGG	CGAATACATC	CCGGCCGTTG	ATAAAGGTAT	1700
	CCAGGAACAG	CTGAAAGCAG	GTCCGCTGGC	AGGCTACCCG	GTAGTAGACA	1750
20	TGGGTATTTCG	TCTGCACTTC	GGTTCTTACC	ATGACGTTGA	CTCCTCTGAA	1800
	CTGGCGTTTA	AACTGGCTGC	TTCTATCGCC	TTTAAAGAAG	GCTTTAAGAA	1850
	AGCGAAACCA	GTTCTGCTTG	AGCCGATCAT	GAAGGTTGAA	GTAGAAACTC	1900
	CGGAAGAGAA	CACCGGTGAC	GTTATCGGTG	ACTTGAGCCG	TCGTCTGTTG	1950
	ATGCTCAAAG	GTCAGGAATC	TGAAGTTACT	GGCGTTAAGA	TCCACGCTGA	2000
25	AGTACCGCTG	TCTGAAATGT	TCGGATACGC	AACTCAGCTG	CGTTCTCTGA	2050
	CCAAAGGTCG	TGCATCATAC	ACTATGGAAT	TCCTGAAGTA	TGATGAAGCG	2100
	CCGAGTAACG	TTGCTCAGGC	CGTAATTGAA	GCCCGTGGTA	AATAA	2145

30

## 2) INFORMATION FOR SEQ ID NO: 1231

## (i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 37 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

40

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1231

GCGAGCCCGA AGATAAAAAA GAACCTCTGC TGCTCGC

37

45

## 2) INFORMATION FOR SEQ ID NO: 1232

## (i) SEQUENCE CHARACTERISTICS:

50

- (A) LENGTH: 42 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

55

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1232

GGAGCCGCGC GATTTTATAA ATGAATGTTG ATAACCGGCT CC

60

42

## 2) INFORMATION FOR SEQ ID NO: 1233

5 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
10 (ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1233  
15 GCGAGCGTTA CTGGTGTAGA AATGTTCCGG CTCGC 35

## 2) INFORMATION FOR SEQ ID NO: 1234

20 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 bases  
(B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
25 (D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:  
30 ACTAAATAAA CGCTCATTCG 20

## 2) INFORMATION FOR SEQ ID NO: 1235

35 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 38 bases  
(B) TYPE: Nucleic acid  
40 (C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1235  
GCGAGCCGAA GTTGAAGTTG TTGGTATTGC TGGCTCGC 38

## 2) INFORMATION FOR SEQ ID NO: 1236

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 bases  
55 (B) TYPE: Nucleic acid  
(C) STRANDEDNESS: Single  
(D) TOPOLOGY: Linear  
(ii) MOLECULE TYPE: DNA  
60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1236

GCGAGCCGTG GTGAAGTTCG CGTTGGTGGC TCGC

34

5

2) INFORMATION FOR SEQ ID NO: 1237

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1237

GCGAGCCGCG AAATCGAAGT TGCTGTATTA GGGCTCGC

38

20

2) INFORMATION FOR SEQ ID NO: 1238

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1238

GCGAGCGGCG TTAATTTTGG CACCGAAGAA GAGCTCGC

38

35

2) INFORMATION FOR SEQ ID NO: 1239

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1239

GCGAGCGCAG ACCTTTCAGC AGAGGAGGCT CGC

33

40

45

50

2) INFORMATION FOR SEQ ID NO: 1240

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single

60

(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: DNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1240

GCGAGCCGGC AAGACAATAT GACAGCAAAA TCGCTCGC

38

10

2) INFORMATION FOR SEQ ID NO: 1241

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

15

(ii) MOLECULE TYPE: DNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1241

GCGAGCGGGG AACGAGGATG ATTTGATTGG CTCGC

35

25

2) INFORMATION FOR SEQ ID NO: 1242

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 600 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

30

(ii) MOLECULE TYPE: Genomic DNA

35

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecium*
- (E) STRAIN: BM4147-1
- (F) ACCESSION NUMBER: U39790

40

(xiv) SEQUENCE DESCRIPTION: SEQ ID NO: 1242

45	TTCTTAGAGA CATTGAATAT GCCTTATGTC GCGCAGGCG TATTGACCAG	50
	TGCATGTGCC ATGGATAAAA TCATGACCAA GTATATTTTA CAAGCTGCTG	100
	GTGTGCCGCA AGTTCCTTAT GTACCAGTAC TTAAGAATCA ATGGAAAAGAA	150
	AATCCTAAAA AAGTATTTGA TCAATGTGAA GGTTCCTTGC TTTATCCGAT	200
	GTTTGTCAAA CCTGCGAATA TGGGTTCTAG TGTCGGCATT ACAAAGGCAG	250
	AAAACCGAGA AGAGCTGCAA AATGCTTTAG CAACAGCCTA TCAGTATGAT	300
50	TCTCGAGCAA TCGTTGAACA AGGAATTGAA GCGCGCGAAA TCGAAGTTGC	350
	TGTATTAGGA AATGAAGATG TTCGGACGAC TTTGCCTGGC GAAGTCGTAA	400
	AAGACGTAGC ATTCTATGAT TATGAAGCCA AATATATCAA TAATAAAATC	450
	GAAATGCAGA TTCCAGCCGA AGTGCCGGAA GAAGTTTATC AAAAAGCGCA	500
	AGAGTACGCG AAGTTAGCTT ACACGATGTT AGGTGGAAGC GGATTGAGCC	550
55	GGTGCGATTT CTTTTTGACA AATAAAAATG AATTATTCCT GAATGAATTA	600

60

2) INFORMATION FOR SEQ ID NO: 1243

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2275 bases  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Double  
 (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Enterococcus faecalis*  
 (C) ACCESSION NUMBER: M38386

## (xv) SEQUENCE DESCRIPTION: SEQ ID NO: 1243

15	GGTACCAAAG	AAAAAAACGA	ACGCCACAAC	CAACAGCCTC	TAAAGCAACA	50
	CCTGCTTCTG	AAATTGAGGG	AGATTTAGCA	AATGTCAATG	AGATTCTTTT	100
	GGTTCACGAT	GATCGTGTCTG	GGTCAGCAAC	GATGGGAATG	AAAGTCTTAG	150
	AAGAAATTTT	AGATAAAGAG	AAAAATTTCAA	TGCCGATTCTG	AAAAATTAAT	200
	ATTAATGAAT	TAACCAACA	AACACAGGCT	TTAATTGTCA	CAAAAGCTGA	250
20	ACTAACGGAA	CAAGCACGTA	AAAAAGCACC	GAAAGCGACA	CACTTATCAG	300
	TAAAAAGTTA	TGGTTAATCC	CCAAAAATAT	GAAACAGTGG	GTTTCGCTCT	350
	TAAAAGAAAG	TGCCTAGAGA	GGAAAGAAAAC	AATGGAAAAT	CTTACGAATA	400
	TTTCAATTGA	ATTAAATCAA	CAGTTTAATA	CAAAAGAAGA	AGCTATTTCGC	450
	TTTTCCGGCC	AGAACTAGT	CGAGGCAGGC	TGTGTTGAGC	CCGCTTATAT	500
25	CGAAGCAATG	ATTGAAAGAG	ACCAATTGCT	ATCTGCCCAT	ATGGGGAATT	550
	TTATTGCCAT	TCCTCATGGA	ACAGAAGAAG	CCAAAAAATT	AGTGAAAAAA	600
	TCAGGAATCT	GTGTAGTGCA	AGTCCCAGAG	GGCGTTAATT	TTGGCACCGA	650
	AGAAGATGAA	AAAATTGCTA	CCGTATTATT	TGGGATTGCC	GGAGTCGGTG	700
	AAGAACATTT	GCAATTAGTC	CAACAAATTG	CACTTTATTG	TAGTGATATG	750
30	GATAACGTGG	TGCAACTTGC	CGATGCATTA	AGTAAAGAAG	AAATAACAGA	800
	AAATTTAGCC	ATTGCTTAAA	GGAGAGAATA	AGAATGAACG	CAGTACATTT	850
	TGGAGCAGGA	AATATTGGAC	GCGGCTTTAT	TGGCGAAATT	TTAGCTAAAA	900
	CGGGTTTCAT	ATTACCGTTT	GTGGATGTTA	ATGGAAACCA	TCATCAAGCG	950
	TTAAAAGAAC	GTAAAAGTTA	TACAATTGAA	TTGGCCGATG	CCTCACATCA	1000
35	ACAAATTAAC	GTTGAAAATG	TGACCGGGTT	AAATAACATG	ACAGAACCAG	1050
	AAAAAGTAGT	AGAAGCAATT	GCGGAAGCCG	ATTTAGTCAC	GACGGCAATT	1100
	GGTCCTAATA	TTTTACCAAG	AATTGCTGAA	TTAATTGCTC	AAGGAATTGA	1150
	TGCACGTGCC	GAAGCAAATT	GTCAAAACGG	CCCGCTGGAT	ATTATCGCTT	1200
	GTGAAAATAT	GATTGGTGGT	TCAACCTTTT	TAGCAGAAGA	AGTGGCCATA	1250
40	ATATTTGAAA	AACCCAGCTT	ATCTGAACAA	TGGATTGGTT	TTCTTGATGC	1300
	GGCAGTTGAT	CGGATTGTTT	CATTACAAAA	ACATAAAGAT	CCACTTTTTG	1350
	TTCAAGTTGA	GCCTTTTTTG	GAATGGGTCA	TTGATGATAC	CAACCGAAAA	1400
	GCCAAAGAGA	TTCAGTTAGA	AGGCGTCATT	ACTTGTCGAT	TAGAGCCGTA	1450
	TATTGAACGA	AAATTATTTA	GTGTAACCAG	TGGCCATGCT	ACAGTTGCCT	1500
45	ATACAGGGGC	GTTGTTAGGC	TATCAAACCA	TTGACGAAGC	GATGCAGGAC	1550
	GCCTTAGTGG	TAGCGCAACT	CAAATCAGTT	TTGCAGGAAA	CCGGTAAACT	1600
	TTTAGTGGCC	AAATGGAATT	TTGATGAACA	AGAACATGCA	GCCTATATTG	1650
	AAAAAATTAT	CAACCGTTTC	CAAAATAAAT	ATATTTTCAGA	TGCTATTACA	1700
	CGTGTAGCAC	GGACACCAAT	CAGAAAATTA	GGTGCGCAAG	AACGGTTTAT	1750
50	TCGACCAATC	CGTGAATTAC	AGGAACGCAA	TCTAGTGTCG	GCCGCATTTA	1800
	TAGCAATGAT	TGGTATTGTC	TTTAATTATC	ATGATCCAGA	AGATGAACAA	1850
	AGCCGTCAAT	TACAGGAAAT	GCTTGACCAA	GAAAGTGTTG	ATACAGTGGA	1900
	TCGCTGAAGT	AACGGGCATT	GAAGATCCAG	AAACGGTTAA	AAATATTAAA	1950
	CAAAACGTAG	AACCTGCTATG	CGCGACCACA	AGTAGCATAA	TTAACAAAAT	2000
55	CCTTCTACCA	AGATACTTCA	CATTTCCTTA	TTAAAGAAAA	AACAACCGCG	2050
	CCTCACCTGA	GCCGACCCCC	AAAAGTTAGA	CCTAGAAATC	TAACCTTTTG	2100
	AGGTTTTTTT	GTATGGCAAA	ATACAGTTTT	GAAATTTAAA	CTTAAACTTG	2150
	TTCATGACTA	CTTATATGGT	CAAGGAGGTC	TAAGGTTTCT	CGCAAAGAAG	2200
	TATGGGTTTA	AAGATAGTCT	CAAATAAGCA	AATGGATAAA	TGCCTATAAA	2250
60	GAACTTGGTG	AAGAAGGGGG	GATCC			2275



## 2) INFORMATION FOR SEQ ID NO: 1244

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 442 bases
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: Double
- (D) TOPOLOGY: Linear

## (ii) MOLECULE TYPE: Genomic DNA

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Staphylococcus aureus* subsp. *aureus*
- (B) STRAIN: ATCC 25923

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1244

20	GATCAATCTT	TGTCGGTACA	CGATATTCTT	CACGACTAAA	TAAACGCTCA	50
	TTCGCGATTT	TATAAATGAA	TGTTGATAAC	AATGTTGTAT	TATCTACTGA	100
	AATCTCATTA	CGTTGCATCG	GAAACATTGT	GTTCTGTATG	TAAAAGCCGT	150
	CTTGATAATC	TTTAGTAGTA	CCGAAGCTGG	TCATACGAGA	GTTATATTTT	200
	CCAGCCAAAA	CGATATTTTT	ATAATCATTA	CGTGAAAAAG	GTTTCCCTTC	250
25	ATTATCACAC	AAATATTTTA	GCTTTTCAGT	TTCTATATCA	ACTGTAGCTT	300
	CTTTATCCAT	ACGTTGAATA	ATTGTACGAT	TCTGACGCAC	CATCTTTTGC	350
	ACACCTTTAA	TGTTATTTGT	TTTAAAAGCA	TGAATAAGTT	TTTCAACACA	400
	ACGATGTGAA	TCTTCTAAGA	AGTCACCGTA	AAATGAAGGA	TC	442

## DEMANDES OU BREVETS VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET  
COMPREND PLUS D'UN TOME.

CECI EST LE TOME 4 DE 4

NOTE: Pour les tomes additionnels, veuillez contacter le Bureau canadien des brevets

## JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE  
THAN ONE VOLUME

THIS IS VOLUME 4 OF 4

NOTE: For additional volumes please contact the Canadian Patent Office